

Psihologia *Fake News*

Acceptarea, distribuirea și
corectarea informațiilor false

Coordonatori:
Rainer Greifeneder, Mariela E. Jaffé,
Eryn J. Newman și Norbert Schwarz



TREI

PSIHOLOGIE
PSIHOTERAPIE

Psihologia *Fake News*

Acceptarea, distribuirea și
corectarea informațiilor false

Coordonatori:
Rainer Greifeneder, Mariela E. Jaffé,
Eryn J. Newman și Norbert Schwarz



TREI

PSIHOLOGIE
PSIHOTERAPIE

Psihologia *Fake News*

Acceptarea, distribuirea și
corectarea informațiilor false

Coordonatori:
Rainer Greifeneder, Mariela E. Jaffé,
Eryn J. Newman și Norbert Schwarz



TREI

PSIHOLOGIE
PSIHOTERAPIE

Psihologia *Fake News*

Acceptarea, distribuirea și
corectarea informațiilor false

Coordonatori:
Rainer Greifeneder, Mariela E. Jaffé,
Eryn J. Newman și Norbert Schwarz



TREI

PSIHOLOGIE
PSIHOTERAPIE

PSIHOLOGIE · PSIHOTERAPIE

Colecție coordonată de
Simona Reghintovschi

VOLUM COORDONAT DE

Rainer Greifeneder, Mariela E. Jaffé,
Eryn J. Newman și Norbert Schwarz

Psihologia *Fake News*

Acceptarea, distribuirea și
corectarea informațiilor false

Traducere din limba engleză de
Vlad Vedeanu



Editori:

Silviu Dragomir

Vasile Dem. Zamfirescu

Director editorial:

Magdalena Mărculescu

Design copertă:

Faber Studio (Adelina Butnaru)

Redactare:

Manuela Sofia Nicolae

Director producție:

Cristian ClauDiu Coban

Dtp:

RĂZVAN NASEA

Corectură:

Andreea-Lavinia Dădârlat

Irina Mușătoiu

Conținutul acestei lucrări electronice este protejat prin copyright (drepturi de autor), iar cartea este destinată exclusiv utilizării ei în scop privat pe dispozitivul de citire pe care a fost descărcată. Orice altă utilizare, incluzând împrumutul sau schimbul, reproducerea integrală sau parțială, multiplicarea, închirierea, punerea la dispoziția publică, inclusiv prin internet sau prin rețele de calculatoare, stocarea permanentă sau temporară pe dispozitive sau sisteme cu posibilitatea recuperării informației, altele decât cele pe care a fost descărcată, revânzarea sub orice formă sau prin orice mijloc, fără consimțământul editorului, sunt interzise. Dreptul de folosință al lucrării nu este transferabil.

Drepturile de autor pentru versiunea electronică în formatele existente ale acestei lucrări aparțin persoanei juridice Editura Trei SRL.

Copyright © 2021 selection and editorial matter, Rainer Greifeneder, Mariela E. Jaffé, Eryn J. Newman, and Norbert Schwarz; individual chapters, the contributors

All Rights Reserved.

Authorised translation from the English language edition published by Routledge, a member of the Taylor & Francis Group.

Titlul original: The Psychology of Fake News

Autori: Rainer Greifender, Mariela E. Jaffé, Eryn J. Newman, and Norbert

Schwarz

Copyright © Editura Trei, 2021, pentru prezenta ediție

O.P. 16, Ghișeul 1, C.P. 0490, București

Tel.: +4 021 300 60 90 ; Fax: +4 0372 25 20 20

e-mail: comenzi@edituratrei.ro

www.edituratrei.ro

ISBN (print): 978-606-40-1008-7

ISBN (epub): 978-606-40-1534-1

Psihologia Fake news

Acest volum examinează fenomenul fake news aducând laolaltă experți renumiți din domenii diferite ale psihologiei și ale altor științe conexe și explorează ceea ce a devenit o trăsătură proeminentă a discursului public de la primul referendum Brexit și de la campania electorală din SUA din 2016.

Informațiile false sunt prezente în numeroase sfere importante ale vieții de zi cu zi, inclusiv în politică, în economie, în medicină și în politicile de sănătate, în jurnalism, în educație și în știință. Într-un climat în care în general adevărul și informațiile false se amestecă și sunt intenționat amestecate, această carte pune întrebarea: Ce anume îi face pe oameni să accepte și să distribuie informații (false) și ce se poate face împotriva informațiilor false? Toate aceste aspecte trebuie înțelese în contextul rețelelor de socializare online, care au schimbat fundamental modul în care informația este produsă, consumată și transmisă. Contribuțiile din acest volum rezumă ultimele descoperiri empirice, ultimele teorii și aplicații și discută idei cu totul noi, precum și direcții de intervenție în viitor împotriva fake news.

Oferind, de asemenea, modalități de abordare a informațiilor false într-o epocă a „faptelor alternative“, această carte este fascinantă și esențială pentru studenți și cercetători în psihologie, comunicare și științe politice și pentru jurnaliști și specialiști în politici publice.

Rainer Greifeneder este profesor de psihologie socială la Universitatea din Basel, Elveția. Cercetările sale se concentrează asupra efectului emoțiilor asupra judecății, experiențelor indivizilor și percepțiilor de excludere socială și asupra modurilor în care indivizii construiesc adevărul.

Mariela E. Jaffé este cercetător postdoctoral în psihologie socială la Universitatea din Basel, Elveția. Cercetările ei se concentrează asupra constructului de adevăr, asupra preferințelor pe care le au indivizii în privința diversității și asupra utilizării de intervenții și instrumente pentru luarea deciziilor.

Eryn J. Newman este lector la Universitatea Națională Australiană. Cercetările ei se concentrează asupra modurilor în care oamenii ajung să aibă amintiri și convingeri despre ce e adevărat și asupra modurilor în care informațiile tangențiale sau „pseudo-dovezile“ pot influența evaluarea noilor informații.

Norbert Schwarz este rector, profesor de psihologie și marketing și co-director al Mind & Society Center din cadrul Universității din California de Sud. Cercetările sale abordează natura dependentă de context și încorporată a judecății și procesului decizional și implicațiile pe care le are în domeniul opiniilor publice, comportamentelor consumatorului și cercetărilor realizate în științe sociale.

Autori

Robert Ackland

Universitatea Australiană Națională, Australia

Dolores Albarracín

Universitatea din Illinois de la Urbana-Champaign, SUA

Jordan R. Axt

Universitatea McGill, Canada

Judit Bar-Ilan

Universitatea Bar-Ilan, Israel

Andrew Dawson

Universitatea din California de Sud, SUA

Rainer Greifeneder

Universitatea din Basel, Elveția

Karl Gwynn

Universitatea Australiană Națională, Australia

Gali Halevi

Facultatea de Medicină Icahn de la Mount Sinai, SUA

Mariela E. Jaffé

University din Basel, Elveția

Madeline Jalbert

Universitatea din California de Sud, SUA

Aaron C. Kay

Universitatea Duke, SUA

Mark J. Landau

Universitatea din Kansas, SUA

Benjamin A. Lyons

Universitatea din Utah, SUA

Elizabeth J. Marsh

Universitatea Duke, SUA

Vittorio Merola

Universitatea Stony Brook, SUA

Eryn J. Newman

Universitatea Australiană Națională, Australia

Daphna Oyserman

Universitatea din California de Sud, SUA

Jason Reifler

Universitatea din Exeter, Marea Britanie

Jon Roozenbeek

Universitatea din Cambridge, Marea Britanie

Norbert Schwarz

Universitatea din California de Sud, SUA

Matthew L. Stanley

Universitatea Duke, SUA

Sander van der Linden

Universitatea din Cambridge, Marea Britanie

Lynn Zhang

Universitatea din California de Sud, SUA

Capitolul 1. Ce e nou și adevărat¹ în privința fake news?

Rainer Greifeneder, Mariela E. Jaffé, Eryn J. Newman și Norbert Schwartz

După Brexit și după campania prezidențială din 2016 din SUA, dicționarul Oxford a ales „post-adevăr“ (post-truth) drept Cuvântul Anului 2016, definit ca „referitor la circumstanțe în care faptele obiective au o influență mai redusă asupra opiniei publice decât apelul la emoții sau la convingerile personale“ (Oxford-Dictionary, 2016). Un an mai târziu, dicționarul Collins a selectat „fake news“ drept Cuvântul Collins al anului 2017, ca reacție la creșterea cu 365% a frecvenței utilizării sale în baza de date Collins Corpus* (Collins-Dictionary, 2017). Pe măsură ce diseminarea fake news a devenit tot mai amplă și a devenit subiect de discurs și preocupare publică, faptul de a numi ceva drept fake news s-a transformat într-un nou instrument din repertoriul propagandei politice. De la președintele SUA Donald Trump, care acuză jurnaliștii și mass-media că răspândesc știri false despre el (e.g. Pengelly, 2017), la mass-media care urmărește știrile false diseminate de Donald Trump (Kessler, Rizzo și Kelly, 2019) și la partidul german AFD (Alternative für Deutschland), care a revenit la termenul nazist de Lügenpresse (presă mincinoasă) pentru a descrie mass-media, acuzațiile de răspândire a știrilor false au devenit un lucru de zi cu zi. Acest climat general reprezintă fundalul pe care volumul de față prezintă analize ale fake news din multiple discipline și perspective științifice.

Ce înseamnă fake news și ce înseamnă informații greșite?

Fake news pot fi definite ca „informații false, adesea senzaționaliste, distribuite sub aparența prezentării de știri“ (Collins-Dictionary, 2017). Analize ale căutărilor pe Google indică faptul că, înainte de 2016, termenul era utilizat pentru a indica știri satirice, ca de exemplu cele din revista The Onion și emisiunea de pamflet The Daily Show. Până la finalul anului 2016, căutările se refereau în mai mare măsură la alegerile prezidențiale din SUA, la Donald Trump, Twitter și CNN (Cunha, Magno, Caetano, Teixeira și Almeida, 2018). Diseminarea de informații false sub aparența prezentării de știri nu mai ținea de amuzament.

Diferența dintre fake news și știrile pur și simplu false este intenția de a induce în eroare. După cum a notat Consiliul European (Wardle și Derakhshan, 2017), discuțiile publice despre fake news includ adesea termenii de dezinformare, informații greșite (misinformation) și informații răuvoitoare (mal-information). Termenul „dezinformare“ se referă la informațiile false create pentru a face rău unei persoane, unui grup social, unei organizații sau unui stat, în timp ce „informațiile greșite“ (misinformation) sunt pur și simplu neadevărate, fără însă a avea intenția de a face rău. Dat fiind că intenția de a face rău este adesea dificil de evaluat, termenul „informații greșite“ (misinformation) este folosit pentru a denota și informațiile false în general și acesta este uzul respectat de contribuțiile la acest volum. Desigur, se poate face rău și prin răspândirea de informații corecte, care au implicații negative — de exemplu, prin publicarea de informații factual adevărate, referitoare la viața privată, care nu ar trebui să afecteze chestiunile publice —, tip de informație denumit uneori „mal-information“. Toate aceste categorii de informații au o influență mai mare când sunt acceptate ca fiind adevărate și sunt transmise și altora. Odată acceptate, informațiile false sunt foarte greu de corectat și pot continua să influențeze convingeri conexe, chiar dacă oamenii nu mai susțin informațiile false care au dus la acele convingeri (pentru o analiză, vezi Lewandowsky, Ecker, Seifert, Schwarz și Cook, 2012). Contribuțiile la acest volum se concentrează asupra proceselor implicate în acceptarea, distribuirea și corectarea informațiilor false.

La nivelul discursului public, termenul fake news este asociat în general cu

sfera politică, dar reportajele fake news nu se limitează la aceasta; în toate domeniile sunt diseminate informații fabricate, poate în mod predominant în privința bunurilor de consum, sănătății și finanțelor. Dar nici revistele științifice respectabile nu sunt lipsite de fake news, sub forma raportărilor intenționat false, bazate pe date manipulate sau pur și simplu inventate (Bar-Ilan și Halevi, 2020).

Fake news au atras atenția publică din mai multe motive. În primul rând, informațiile false au ajuns să facă parte din viața de zi cu zi (vezi Lyons, Merola și Reifler, 2020). De exemplu, Allcott și Gentzkow (2017) au estimat că un adult obișnuit din SUA a consumat una până la trei știri false în lunile de dinaintea alegerilor prezidențiale din 2016 și jurnaliștii de la Washington Post care se ocupă de verificarea acurateții informațiilor (fact checkers) au raportat că Trump a făcut 10 796 de afirmații false sau inexacte în primele 869 de zile de la câștigarea alegerilor (Kessler et al., 2019). Mai mult, Vosoughi, Roy și Aral (2018) au descoperit prin analize ale datelor Twitter că informațiile false sunt redistribuite mai repede și mai des decât informațiile adevărate, în special știrile politice.

În al doilea rând, în măsura în care oamenii cred informațiile false și acționează pe baza acestora, fake news pot avea consecințe serioase. De pildă, în timpul a ceea ce a ajuns să fie numit Pizzagate (Wikipedia, nedatat), o știre falsă despre o presupusă legătură între oficiali ai Partidului Democrat, o pizzerie din Washington, D.C. și un lanț de traficanți de oameni l-a făcut pe un tânăr să „investigheze” el însuși problema și a tras cu arma în restaurant. Nu toate consecințele directe ale știrilor false sunt la fel de dramatice; dar, când indivizi sau grupuri acționează pe baza dezinformării, consecințele pot adesea să nu fie în avantajul indivizilor sau societății în general.

În al treilea rând, interacțiunile umane pașnice și prosperitatea individuală și socială depind în mare măsură de încrederea interpersonală (e.g. Greifeneder, Müller, Stahlberg, Van den Bos și Bless, 2011). Diviziunea muncii, comerțul între persoane și state, precum și democrațiile, toate necesită credințe împărtășite că unele lucruri sunt adevărate și că agenții sunt de încredere. Indivizii găsiți vinovați de a fi mințit nu sunt (sau sunt mai puțin) crezuți, iar companiile și statele despre care se știe că au încălcat

promisiuni financiare se confruntă cu consecințe serioase. Știrile false despre un guvern au potențialul de a eroda încrederea societății și constituie astfel o amenințare, în special la adresa democrațiilor. Mai mult, denunțarea ziarelor, posturilor de radio și revistelor consacrate are potențialul de a eroda încrederea în aceste organizații care își asumă rolul de a verifica faptele în societățile moderne. În SUA, nivelul de încredere în sursele mass-media de știri a relevat o diviziune pronunțată unilaterală în perioada alegerilor din 2016, 51% dintre democrați și doar 14% dintre republicani considerând că au „destul de multă” sau „foarte multă” încredere în sursele mass-media de știri (Swift, 2016).

Scurtă istorie a diseminării fake news

Intenția de a induce în eroare e la fel de veche ca omenirea și campanii sistematice de fake news au fost documentate pe întreg parcursul istoriei (vezi Posetti și Matthews, 2018). Ce s-a schimbat este capacitatea de a răspândi informații false rapid și eficient în rândul unor audiențe din ce în ce mai mari. Când Ramses al II-lea s-a prefăcut în 1274 î.H. că a reușit în încercarea lui de a cuceri orașul Kadesch, vestea falsei victorii a trebuit să fie răspândită prin viu grai și prin desene murale. Când Octavian și-a dus campania de propagandă împotriva lui Antoniu în Roma antică, a putut să-l înfățișeze ca afemeiat și bețiv prin sloganuri înscrise pe monede, ceea ce a permis o distribuie amplă (Kaminska, 2017). Inventarea de către Gutenberg a tiparului în 1493 a făcut posibilă comunicarea la scară largă către multe persoane și, astfel, diseminarea și de informații adevărate, și de informații false. În 1835, un ziar pe atunci tipărit la New York, The Sun, a publicat o serie de articole despre descoperirea de creaturi umanoide pe Lună, ceea ce a ajuns să fie cunoscut ca Marea Farsă Lunară (Thornton, 2000). Apariția radioului a facilitat și mai mult diseminarea știrilor de toate felurile, inclusiv purtarea campaniilor masive de dezinformare în perioada de dinaintea celui de-al Doilea Război Mondial (Herzstein, 1978; Kallis, 2005).

Cu toate astea, utilizarea sistematică de tipărituri și emisiuni radio-TV și diseminarea produselor acestora la scară largă necesita resurse considerabile, ceea ce a limitat numărul actorilor care puteau profita de aceste tehnologii. Aspectul negativ e că asta le permitea actorilor care aveau acces la ele să răspândească informația falsă fără opoziție semnificativă; aspectul pozitiv e că astfel a apărut și șansa de a modela transmiterea informațiilor prin implementarea unor norme etice ale jurnalismului (Ward, 2015; vezi și contribuțiile din Bertrand, 2018). Apariția internetului, urmată de dezvoltarea rețelelor de socializare, a redus barierele existente, astfel că acum majoritatea indivizilor pot să participe la diseminarea pe scară largă. Prin tehnologia modernă a telefoanelor inteligente și prin rețelele de socializare online, orice utilizator de internet poate fi o sursă de informații. De pildă, utilizatorul obișnuit de Twitter are 707 de follower-i (Smith, 2019) către care informația poate fi transmisă instantaneu. Acest nivel de acces era înainte imposibil pentru indivizi, ceea ce face ca această eră să fie fundamental nouă — diseminarea de știri nu mai este un privilegiu al câtorva, ci este disponibilă pentru toți (utilizatorii de internet). Deși această schimbare are potențialul de a le da putere cetățenilor, permite și răspândirea necontrolată de informații false și impune dezvoltarea de noi norme sociale de evaluare și distribuire a informațiilor.

Astăzi, site-urile de social media decid prin intermediul algoritmilor ce știri vor vedea utilizatorii lor. În speță, companii ca Facebook filtrează știrile disponibile și le prezintă utilizatorilor o selecție. Detaliile algoritmului le sunt necunoscute utilizatorilor și trec prin schimbări frecvente. Ce se știe este că algoritmul favorizează informațiile care se potrivesc profilului de preferințe ale utilizatorului și exclude informațiile care nu corespund acestuia. Bula filtrantă care rezultă astfel (Pariser, 2011) prezintă informații în mare consecvență, care confirmă perspectiva utilizatorului și prezintă puține provocări, susținând astfel convingerea utilizatorilor că propriile lor perspective sunt corecte, iar perspectivele celorlalți sunt, în cel mai bun caz, greșite, și în cel mai rău caz, răuvoitoare (pentru o discuție, vezi Schwarz și Jalbert, 2020). Numeroși observatori bănuiesc că astfel de bule filtrante au contribuit la rezultatul votului Brexit din 2016 (vezi Oyserman și Dawson, 2020). În combinație cu omofilia naturală a rețelelor de socializare, care face ca indivizii să se împrietenească de obicei cu alții care gândesc asemănător, mecanismele de filtrare pot crea rețele omogene puternice, în care conținutul

care corespunde valorilor și normelor membrilor are o probabilitate mai mare de a fi comunicat. Un aspect esențial este că informația difuzată în astfel de bule poate ajunge să sune ca un ecou, permițându-se astfel unor voci izolate să pară un cor.

Eficienței de diseminare pe care o au rețelele de social media i se adaugă și faptul că nu mai e necesar ca agenții jocului de distribuire a informațiilor să fie ființe umane. Boții sociali pot să genereze, să distribuie, să redistribuie și să dea like fără să fie ghidate deloc sau aproape deloc de oameni. Asta afectează conținutul, volumul și valoarea dată (îmi place/nu-mi place) informației și poate modifica semnificativ mediul informațional în anumite bule. Deși Twitter, Facebook și alte rețele încearcă să reducă numărul de conturi automate, e vorba de o cursă a înarmării în care cei care doresc să inducă în eroare își adaptează comportamentul pentru a ocoli sau a depăși ultima tehnologie dezvoltată de cei care doresc să restricționeze rețeaua la agenții umani. În viitorul foarte apropiat, perfecționarea conținuturilor deep fake — înregistrări video sau fotografii extrem de dificil de identificat ca fiind fabricații înșelătoare — va aduce și mai multe provocări menținerii unui mediu informațional credibil.

Volumul de față

Într-un climat în care faptele și informațiile false se amestecă și sunt intenționat amestecate, acest volum întreabă ce anume îi face pe oameni să accepte și să distribuie informații (sau informații false) și ce se poate face pentru a contracara informațiile false. Toate cele trei aspecte trebuie înțelese în contextul rețelilor de socializare online, care au schimbat în mod fundamental modul în care informația este produsă, consumată și transmisă. Pentru a răspunde acestor întrebări, volumul de față aduce laolaltă experți consacrați din diferite domenii ale psihologiei și ale științelor înrudite, precum științele informației și științele politice.

Acest volum este împărțit în trei secțiuni. Prima secțiune se concentrează asupra originilor și urmărilor fake news în rețelele de socializare online și în mediul academic. Lyons, Merola și Reifler (2020) întreabă „Cât de gravă e problema fake news?” și prezintă date despre consumul perceput de fake news după campania prezidențială din 2016 din SUA. Ackland și Gwynn (2020) oferă un raport bazat pe cifre asupra difuzării de știri pe Twitter ca funcție a statutului de adevăr, adică dacă știrile erau adevărate sau false. Autorii oferă și o evaluare de ansamblu asupra literaturii despre verificarea faptelor (fact checking). În sfârșit, Bar-Ilan și Halevi (2020) investighează consecințele contribuțiilor de cercetare care au fost retractate, în special cele care au provenit din încălcarea deontologiei, inclusiv din fabricarea de date și falsificări. Ei se concentrează asupra exemplelor din științe medicale și biomedicale, dat fiind efectul lor potențial dezastruos asupra sănătății publice.

Secțiunea a doua și a treia ale volumului sunt constituite în principal din cercetări psihologice. Comunicarea servește necesităților informaționale și sociale pe care oamenii le au în mod fundamental. Ingeniozitatea rețelelor de socializare online stă în faptul că răspund acestor nevoi, permițându-le oamenilor să transmită informația, să învețe, să fie surprinși, informați și să se afle în postura de a evalua conținutul și de a-i evalua pe ceilalți. Însă în satisfacerea acestor nevoi multe procese motivaționale și cognitive pot acționa ele însele ca filtre și distorsionări, afectând tipul de informații pe care indivizii le caută, le percep, le cred adevărate sau le actualizează în mod selectiv. Rămâne o provocare enormă în ce privește combaterea știrilor false faptul că oamenii tind să fie mai creduli când e vorba de lucruri care își doresc să fie adevărate sau pe care le cred adevărate (despre puterea distorsiunilor de confirmare, e.g. Nickerson, 1998).

În centrul celei de-a doua secțiuni a volumului se află procesele cognitive implicate în acceptarea, redistribuirea și corectarea informațiilor false. Schwarz și Jalbert (2020) analizează principalele criterii pe care oamenii le folosesc pentru a stabili dacă ceva e probabil să fie adevărat. Ei subliniază faptul că mesajul are cea mai mare probabilitate de a fi acceptat dacă este compatibil cu alte lucruri pe care persoana le cunoaște, dacă are coerență internă, dacă este prezentat de o sursă credibilă, dacă este acceptat de alții similari și dacă este susținut de dovezi. Fiecare dintre aceste criterii poate fi evaluat pe baza unor informații considerabile; însă criteriile pot fi evaluate și

prin investigarea experienței metacognitive a fluenței, adică a gradului de dificultate cu care mesajul poate fi procesat. Bazându-se pe cercetări experimentale, Schwartz și Jalbert arată cum operează fluența și discută implicațiile acestora în contextul acceptării fake news.

O imagine valorează cât o mie de cuvinte. Influențează și ce credem, surprinzător, chiar dacă imaginea nu conține (suplimentar) informații care să confirme asta, după cum indică Newman și Zhang (2020). Autorii analizează cercetări experimentale asupra puterii imaginilor de a da impresia de adevăr și discută implicațiile pentru constituirea credințelor și pentru memorie. Analizează și examinează riguros și motivul pentru care imaginile au puterea de a induce în eroare și oferă concluzii referitoare la condițiile minime pe care trebuie să le îndeplinească fotografiile pentru a produce un efect de distorsionare.

Trecând de la imagini la structurarea frazelor, Jaffé și Greifeneder (2020) arată că până și simpla încadrare conceptuală a informației în termeni negativi poate determina un nivel mai ridicat de încredere. Autorii dezvoltă în continuare un cadru al antecedentelor acestui efect, arătând, de exemplu, că așteptările joacă un rol vital.

Marsh și Stanley (2020) aduc în centrul atenției procesele memoriei, în special construirea, reprezentarea și actualizarea cunoștințelor. Ghidați de întrebarea dacă convingerile eronate sunt un produs firesc al unui sistem adaptativ de cunoaștere, ei oferă idei bazate pe științele cognitive în privința strategiilor de contracarare a fake news care au șansa de a reuși sau care probabil vor eșua.

Indivizii asociază în general inocularea cu tratamentul medical. Dar conceptul de inoculare a fost aplicat și în cazul convingerilor, presupunerea de bază fiind că respingerea unui atac minor poate amplifica probabilitatea de contracarare a unor atacuri ulterioare la scară mai mare (McGuire, 1964). Aplicând această noțiune în cazul fake news, van der Linden și Roozenbeek (2020) indică modalități prin care un „vaccin psihologic“ sub forma educației mass-media poate pregăti indivizii să contracareze fake news.

În a treia secțiune a volumului, cercetătorii își îndreaptă atenția asupra proceselor motivaționale. Oyserman și Dawson (2020) oferă dovezi

consistente că motivația bazată pe identitate modelează ceea ce oamenii cred, distribuie și acceptă. Abordarea lor subliniază unul dintre pilonii fundamentali ai psihologiei sociale: indivizii nu reacționează la mediul obiectiv, ci la constructul său subiectiv, influențat semnificativ de aspirațiile indivizilor și de modul în care aceștia se văd pe ei înșiși.

Procesele motivaționale sunt în centrul ideilor oferite și de Albarracín (2020), care analizează cercetările asupra convingerilor conspiraționiste. Ea explică procesele care amplifică aceste convingeri și subliniază rolul motivațiilor cu rol de apărare a eului.

În sfârșit, Axt, Landau și Kay (2020) se concentrează asupra aspectului propagandistic al fake news, a cărui intenție este discreditarea surselor tradiționale de știri, fapt cel mai bine ilustrat de termenul nazist Lügenpresse (presă mincinoasă) reintrodus de partidul german AFD și reflectat în referințele lui Donald Trump la „falsele” și „mincinoasele” mass-media. Autorii discută faptul că noțiunea de presă care induce intenționat în eroare poate fi deosebit de atrăgătoare pentru indivizii care au o mare nevoie de structură, dat fiind că o lume în care presa are intenția de a înșela este mai structurată decât o lume în care presa este pur și simplu neglijentă.

Notă

1.Din câte știm astăzi.

Referințe

Ackland, R. și Gwynn, K. (2020). Truth and the dynamics of news diffusion on twitter. În R. Greifeneder, M. Jaffé, E.J. Newman și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 27–46). Londra: Routledge.

Albarracín, D. (2020). Conspiracy beliefs: Knowledge, ego-defense, and social integration. În R. Greifeneder, M. Jaffé, E.J. Newman și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 196–219). Londra: Routledge.

Allcott, H. și Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31, 211–236. doi: 10.1257/jep.31.2.211

Axt, J.R., Landau, M.J. și Kay, A.C. (2020). Fake news attributions as a source of nonspecific structure. În R. Greifeneder, M. Jaffé, E.J. Newman și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 220–234). Londra: Routledge.

Bar-Ilan, J. și Halevi, G. (2020). Retracted articles: The scientific version of fake news. În R. Greifeneder, M. Jaffé, E.J. Newman și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 47–70). Londra: Routledge.

Bertrand, C.-J. (2018). *Media ethics and accountability systems*. New York: Routledge.

Collins-Dictionary. (2017). Collins 2017 word of the year shortlist. Accesat la 2 noiembrie 2019, www.collinsdictionary.com/word-lovers-blog/new/collins-2017-word-of-the-year-shortlist,396, HCB.html

Cunha, E., Magno, G., Caetano, J., Teixeira, D. și Almeida, V. (2018). Fake news as we feel it: Perception and conceptualization of the term „fake news“ in the media. În S. Staab, O. Koltsova și D. Ignatov (coord.), *Social informatics: SocInfo 2018: Lecture notes in computer science* (Vol. 11185, pp. 151–166). Heidelberg: Springer.

Greifeneder, R., Müller, P., Stahlberg, D., Van den Bos, K. și Bless, H.

(2011). Guiding trustful behavior: The role of accessible content and accessibility experiences. *Journal of Behavioral Decision Making*, 24, 498–514. doi: 10.1002/bdm.705

Herzstein, R.E. (1978). *The war that Hitler won: The most infamous propaganda campaign in history* (Vol. 1977). New York: Putnam Publishing Group.

Jaffé, M.E. și Greifeneder, R. (2020). Can that be true or is it fake news? New perspectives on the negativity bias in judgments of truth. În R. Greifeneder, M. Jaffé, E.J. Newman, și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 115–130). Londra: Routledge.

Kallis, A. (2005). *Nazi propaganda and the second world war*. New York: Palgrave Macmillan.

Kaminska, I. (2017). A lesson in fake news from the info-wars of ancient Rome. *Financial Times*. Accesat la <https://www.ft.com/content/aaf2bb08-dca2-11e6-86acf253db7791c6>

Kessler, G., Rizzo, S. și Kelly, M. (2019). President Trump has made 10,796 false or misleading claims over 869 days. *Washington Post*. Accesat la <https://www.washingtonpost.com/politics/2019/06/10/president-trump-has-made-false-or-misleading-claims-over-days/>

Lewandowsky, S., Ecker, U.K.H., Seifert, C.M., Schwarz, N. și Cook, J. (2012). Misinformation and its correction continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13, 106–131. doi: 10.1177/1529100612451018

Lyons, B., Merola, V., și Reifler, J. (2020). How bad is the fake news problem? The role of baseline information in public perceptions. În R. Greifeneder, M. Jaffé, E.J. Newman, și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 11–26). Londra: Routledge.

Marsh, E.J., și Stanley, M. (2020). *False beliefs: Byproducts of an adaptive*

knowledge base? În R. Greifeneder, M. Jaffé, E.J. Newman și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 131–146). Londra: Routledge.

McGuire, W.J. (1964). Inducing resistance to persuasion: Some contemporary approaches. *Advances in Experimental Social Psychology*, 1, 191–229.

Newman, E.J. și Zhang, L. (2020). Truthiness: How non-probative photos shape belief. În R. Greifeneder, M. Jaffé, E.J. Newman și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 90–114). Londra: Routledge.

Nickerson, R.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2, 175. doi: 10.1037/1089–2680.2.2.175

Oxford-Dictionary. (2016). Word of the year 2016 is... Accesat la 2 noiembrie 2019, <https://languages.oup.com/word-of-the-year-2016###targetText=After%20much%20discussion%2C%20debate%2C%20an>

Oyserman, D. și Dawson, A. (2020). Your fake news, our fakes: Identity-based motivation shapes what we believe, share, and accept. În R. Greifeneder, M. Jaffé, E.J. Newman și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 173–195). Londra: Routledge.

Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. New York: Penguin.

Pengelly, M. (2017). Trump accuses CNN of „fake news“ over reported celebrity apprentice plans. Accesat la 30 iulie 2019, www.theguardian.com/us-news/2016/dec/10/trump-celebrity-apprentice-cnn-fake-news

Posetti, J. și Matthews, A. (2018). A short guide to the history of „fake news“ and disinformation: A new ICFJ learning module. Accesat la 2 noiembrie 2019, www.icfj.org/news/short-guide-history-fake-news-and-disinformation-

new-icfj-learning-module

Schwarz, N. și Jalbert, M. (2020). When news feels true: Intuitions of truth and the acceptance and correction of misinformation. În R. Greifeneder, M. Jaffé, E.J. Newman și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 73–89). Londra: Routledge.

Smith, K. (2019). 58 incredible and interesting twitter stats and statistics. Accesat la 31 iulie 2019, www.brandwatch.com/blog/twitter-stats-and-statistics/

Swift, A. (2016). Americans' trust in mass media sinks to new low. Accesat la 2 noiembrie 2019, <https://news.gallup.com/poll/195542/americantrust-mass-media-sinks-new-low.aspx>

Thornton, B. (2000). The moon hoax: Debates about ethics in 1835 New York newspapers. *Journal of Mass Media Ethics*, 15, 89–100. doi: 10.1207/S15327728JMME1502_3

van der Linden, S., și Roozenbeek, J. (2020). A psychological vaccine against fake news. În R. Greifeneder, M. Jaffé, E.J. Newman și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 147–169). Londra: Routledge.

Vosoughi, S., Roy, D. și Aral, S. (2018). The spread of true and false news online. *Science*, 359, 1146–1151. doi: 10.1126/science.aap9559

Ward, S.J.A. (2015). *The invention of journalism ethics: The path to objectivity and beyond*. Montreal: McGill-Queen's Press.

Wardle, C. și Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. Strasbourg: Council of Europe. Wikipedia. (nedatat). Pizzagate conspiracy theory. Accesat la 30 iulie 2019, https://en.wikipedia.org/wiki/Pizzagate_conspiracy_theory.

* Bază de date actualizată constant, creată în scopul analizei lingvistice și care cuprinde utilizările comune ale cuvintelor englezești preluate din cărți, ziare, reviste, transmisiuni radio și TV, materiale online și din conversații cotidiene. (N.t.)

PARTEA I. Parcursul și consecințele informației (false) în rețele

Capitolul 2. Cât de gravă este problema fake news?

Rolul bazelor de referință asupra percepțiilor publicului

Benjamin A. Lyons, Vittorio Merola și Jason Reifler

În februarie 2019, parlamentul britanic a lansat un raport dur în care asemena conducerea Facebook cu niște „gangsteri digitali“, din cauza modului în care tratau datele utilizatorilor. Acest raport sever făcea parte din efortul mai amplu de a înțelege rolul specific pe care social media îl joacă în alegeri și rolul pe care îl au în general în subminarea instituțiilor democratice. De o importanță majoră era modul în care platformele de social media contribuie la răspândirea de „fake news“ și de alte forme de dezinformare. E de înțeles de ce parlamentul britanic s-a arătat interesat de modul în care mediile de comunicare digitale pot afecta negativ democrația. La urma urmei, firma britanică Cambridge Analytica a fost amestecată în scandaluri pentru rolul pe care l-ar fi jucat în două alegeri cu rezultate surprinzătoare — referendumul Brexit și victoria șocantă a lui Donald Trump în alegerile prezidențiale din 2016.

În timpul acestor campanii — în special alegerile prezidențiale din SUA — noua formă de conținut politic „fake news“ a început să primească atenție. Fake news produceau conținut fals și care nu ținea deloc seama de acuratețe sau de standarde jurnalistice. Aceste site-uri se prezentau adesea ca site-uri

autentice de știri. Unele dintre cele mai infame spuneau că Papa Francis îl susținea pe Trump (nu a făcut asta), că un agent FBI implicat în publicarea e-mailurilor lui Clinton a fost implicat într-un caz de crimă urmată de sinucidere (asta nu s-a întâmplat) și că cei care protestau față de rezultatul alegerilor erau plătiți de George Soros (nici asta nu era adevărat). Unele dintre aceste articole au fost distribuite pe Facebook de sute de mii — sau de milioane — de ori. O nouă formă de media, al cărei nivel de implicare e scandalos de mare, la care se adaugă apariția simultană a unor rezultate neașteptate la alegeri, reprezintă o combinație care a dus la o inferență firească, chiar dacă naivă — mulți oameni consumau fake news, iar acest consum a avut un efect asupra rezultatului alegerilor.

Pe de altă parte, difuzarea de știri anecdotice este înlocuită acum cu difuzarea de cercetări empirice ale fake news (Grinberg, Joseph, Friedland, Swire-Thompson și Lazer, 2019; Guess, Nyhan și Reifler, 2018; Guess, Nagler și Tucker, 2019). Contrar disperării ce transpare din descrierile populare, fake news au reprezentat un consum limitat de știri în timpul alegerilor din 2016. În speță, acest studiu a descoperit că aproximativ unul din patru (27%) adulți americani au vizitat un website de fake news în perioada anterioară alegerilor din 2016, în medie fiecare vizitând 5,5 articole, iar „website-urile de fake news au reprezentat o medie de aproximativ 2,6% din totalul articolelor de știri pe care americanii le-au citit pe site-uri” în acea perioadă (Guess et al., 2018). Expunerea la fake news a scăzut și mai mult între 2016 și 2018 (Guess et al., 2019).

Titlurile de știri pot simplifica aceste descoperiri, oferindu-le cititorilor un rezumat contextualizat, cum ar fi „Fake News: extindere mare, impact minor, conform oamenilor de știință” (Carey, 2018) sau „Majoritatea americanilor nu au fost expuși la «Fake news» în alegerile din 2016, conform unui studiu asupra Twitter” (Fox, 2019). Cititorii sunt influențați semnificativ de titlul unui articol (e.g. Ecker, Lewandowsky, Chang și Pillai, 2014). Cu toate acestea, e mai puțin clar modul în care bazele de referință statistice stabilite în cercetări pot ele însele să influențeze perspectiva publicului asupra fake news (e.g. Kessler, Levine, Opitz și Reynolds, 1987) — în general, în absența unui titlu care rezumă, cum înțelege publicul informația referitoare la numărul de persoane care au fost expuse la fake news și la numărul de articole citite?

Teoria

Efecte potențiale ale bazelor de referință

Cercetările arată că expunerea la fake news în timpul alegerilor din 2016 a fost mai redusă decât pretinde imaginația populară și a fost limitată la un segment mic de populație (Grinberg et al., 2019; Guess et al., 2018; Guess, Lyons, Montgomery, Nyhan și Reifler, 2019). Totuși, această cercetare științifică are o probabilitate redusă de a captiva publicul la fel de puternic pe cât a făcut-o introducerea inițială a problemei fake news. Publicul probabil nu cunoaște prevalența reală a consumului de fake news și, astfel, face erori de inferență (e.g. Ahler și Sood, 2018) pe baza cazurilor care ies în evidență. Astfel, sondajele de opinie arată în general un nivel ridicat de preocupare. Pew Research Center (Barthel, Mitchell și Holcomb, 2016), de exemplu, a descoperit că 64% dintre americani credeau că fake news au stârnit „o mare confuzie” după alegeri și 71% au spus că au văzut frecvent sau uneori știri politice „complet inventate” pe internet.

Oferirea unei baze de referință statistice a informațiilor referitoare la consum poate influența judecățile subiective în privința prevalenței consumului de fake news, a prevalenței în timp, a prevalenței în cadrul anumitor subgrupuri demografice și a interesului general pentru această problemă și pentru găsirea de soluții. Însă este neclar în ce direcție pot aceste baze de referință să influențeze percepțiile și atitudinile. Este posibil ca informația să fie opusă unei presupuse prevalențe larg răspândite și să determine reducerea impresiei subiective a prevalenței și interesului. Pe de altă parte, publicului îi poate fi dificil să pună bazele de referință într-un context adecvat, iar prezentarea informației despre consum, indiferent de volum, poate funcționa ca indicator care determină ieșirea în evidență a problemei fake news, ceea ce ar duce la o

impresie subiectivă mai accentuată în privința prevalenței și interesului. Efectele informației statistice pot depinde și de caracteristici individuale, cum ar fi familiaritatea (e.g. utilizarea de Facebook), cunoștințele politice și interesul față de politică, precum și de trăsături cognitive ca tendința la reflecție, necesitatea de a avea dovezi și încrederea în intuiție.

Reducerea prevalenței percepute și a interesului

O modalitate prin care bazele de referință statistice ale consumului de fake news pot determina reducerea consumului perceput și a interesului constă în oferirea unei norme descriptive (e.g. Rimal și Real, 2005). Pe baza datelor despre opinia publică (Barthel et al., 2016), percepțiile din prezent ale publicului referitoare la consumul de fake news pot fi văzute ca o eroare normativă de percepție (Neighbors, Dillard, Lewis, Bergstrom și Neil, 2006). Astfel, oferirea de informații de referință reale poate alinia percepția cu realitatea. Bazele de referință statistice ale consumului real din timpul alegerilor din 2016 — 27% dintre americani au fost expuși la cel puțin un articol fake news, cu o medie de 5,5 articole pentru fiecare — pot fi opuse unor inferențe extreme, derivate din experiența anecdotică, și pot servi la reducerea nivelului evaluat subiectiv al consumului public (e.g. „Foarte puțin?“ „Mult?“) și, de asemenea, la reducerea interesului constant pentru această chestiune.

Amplificarea prevalenței percepute și a interesului

Pe de altă parte, este renumită nepriceperea oamenilor de a înțelege informații statistice (Hoffrage, Lindsey, Hertwig și Gigerenzer, 2000). Oferirea de baze de referință poate astfel să determine creșterea nivelului perceput subiectiv al extinderii fake news și a interesului personal pentru această chestiune. Este important faptul că „nu detaliile obiective, ci semnificația personală a informației este stocată și reactualizată în luarea de decizii” (Kessler et al., 1987, p. 367; Neisser, 1978). Orice bază de referință nu poate servi decât la concretizarea unei probleme care până atunci era vagă.

Euristica disponibilității (e.g. An, 2008; Tversky și Kahneman, 1974) poate explica un astfel de efect. Ușurința cu care indivizii își amintesc exemple de evenimente influențează percepțiile și convingerile ulterioare referitoare la probabilitatea sau prevalența unor astfel de evenimente (Folkes, 1988; MacLeod și Campbell, 1992). Oferirea de baze de referință statistice poate servi ca un fel de exemplu, ceea ce ar face ca problema consumului de fake news să fie în mai mare măsură prezentă în mintea cititorului, evaluarea subiectivă a prevalenței și interesului fiind astfel mai accentuată. Cu alte cuvinte, prin scoaterea în prim-plan a fake news, baza de referință a informațiilor despre consum poate determina amplificarea intensității evaluărilor subiective (Carroll, 1978).

Efecte asupra distorsionărilor grupo-centrice

Pe lângă supraestimarea generalizată a consumului de fake news (Barthel et al., 2016), este probabil ca indivizii să prezinte distorsiuni grupo-centrice în privința convingerilor referitoare la persoanele care consumă fake news (e.g. Ahler și Sood, 2018; Turner, Brown și Tajfel, 1979). Votanții care o susțineau pe Hillary Clinton susțineau probabil că partizanii lui Trump consumau fake news în mai mare măsură decât susținătorii lui Clinto, și viceversa. Votanții de peste 69 de ani presupun poate că votanții mai tineri au un consum mai mare de fake news decât cei din propria grupă de vârstă și

invers. Absolvenții de facultate pot presupune că persoanele fără studii superioare sunt mai susceptibile la efectul fake news, deși nu e clar dacă această relație ar fi simetrică, date fiind presupunerile răspândite despre efectele pozitive ale educației superioare. Oferirea unei baze de referință a informației în privința consumului general de fake news ar putea accentua aceste diferențe în cazul în care participanților li se cere să evalueze dacă diferite grupuri demografice consumă fake news mai puțin decât media, mai mult decât media sau la fel ca media. Dacă tratamentul bazei de referință statistice determină ieșirea în evidență a fake news și accentuarea evaluării subiective a prevalenței acestora, se poate ca indivizii să proiecteze apoi distorsiuni favorabile asupra distribuției evaluărilor lor amplificate ale consumului general. Dacă în mintea cititorului fake news reprezintă o problemă mai mare, atunci se poate să explice acest nivel mai ridicat de consum perceput prin atribuirea sa (disproporționată) unui alt grup decât categoria socială din care face parte (Turner et al., 1979).

Efecte condiționale ale bazelor de referință

Familiaritatea cu platformele social media și cunoștințele generale de politică îi pot face pe indivizi să treacă cu vederea noi informații despre consumul de fake news din timpul alegerilor din 2016. Date fiind cunoștințele lor ample despre această chestiune, cei care folosesc din plin social media și cei care se pricep la politică ar trebui să aibă o probabilitate mai mică de a-și schimba opiniile referitoare la fake news în fața descoperirilor noi prezentate pe scurt (e.g. Hill, 2017). Invers, cei care până atunci se vor fi gândit mai puțin la asta s-ar putea să fie în mai mare măsură influențați de baza de referință a informației. În mod similar, diferențele dintre trăsăturile cognitive care determină modalitatea în care indivizii interacționează cu informații noi — cum ar fi reflecția cognitivă (Frederick, 2005), necesitatea de a avea dovezi (Garrett și Weeks, 2017) și încrederea în intuiție (Garrett și Weeks, 2017) — îi pot face pe indivizi să fie mai mult sau mai puțin deschiși la noi informații statistice despre consumul de fake news în privința formării de evaluări

subiective. Cei care au o tendință mai accentuată de a reflecta, cei care au o nevoie mai mare de dovezi pentru formarea convingerilor și cei care se bazează mai puțin pe intuiție se poate să aibă o probabilitate mai mare de a fi influențați de noi date.

Efecte ale diferitelor baze de referință

Diferite forme de informații statistice, cum ar fi mediile și procentele, pot fi percepute diferit (Garcia-Retamero și Galesic, 2009; Schapira, Davids, McAuliffe și Nattinger, 2004; Westwood, Messing și Lelkes, 2019). Acesta este motivul pentru care aplicăm abordări diferite. Testăm efectele informației referitoare atât la procentul populației expuse la orice tip de fake news (aproximativ 27% în perioada anterioară alegerilor din 2016), cât și la numărul mediu de articole fake news pe care fiecare american le-a citit în acea perioadă (5,5 articole). Testăm, de asemenea, un tratament** care include ambele baze de referință. Combinarea mai multor baze de referință poate servi ca indicator de amplificare a semnalului, ducând astfel la scoaterea în evidență a problemei fake news.

Metode

Eșantion

Datele noastre provin dintr-un eșantion de 981 de adulți americani recrutați

prin Amazon Mechanical Turk*** în ianuarie 2019. Participanții au primit 0,55 de dolari pentru răspunsurile pe care le-au oferit în medie în 5,92 minute (abatere standard de 6,42). Vârsta medie a eșantionului este de 36,57 ani (abatere standard de 11,03). Puțin mai mult de jumătate din eșantion a fost alcătuită din bărbați (55%) și aproximativ trei sferturi albi (76%). Mediana respondenților a fost alcătuită din absolvenți de facultate (46% nu au educație universitară).

Am evaluat partizanatul politic utilizând formatul standard American National Election Study, în care mai întâi li se cere respondenților să indice cu ce partid se identifică. Celor cărora răspund că sunt republicani sau democrați le este adresată o întrebare prin care se evaluează gradul aderării lor (fie „puternică“, fie „slabă“). Cei care inițial nu se identifică cu un partid le este adresată o altă întrebare despre partidul spre care „tind“; pe aceștia i-am considerat partizani. (Deși aceste evaluări sunt adesea utilizate pentru construirea unei scale cu șapte puncte, de la „susținere intensă a Partidului Democrat“ la „susținere intensă a Partidului Republican“, noi am creat variabile binare pentru aderența la Partidul Democrat și la Partidul Republican). Puțin mai mult de jumătate s-au identificat ca democrați (56%) și aproximativ o treime ca republicani (33%), restul neidentificându-se cu niciun partid. Ideologia a fost aproape identică (51% s-au autoevaluat ca liberali și 30%, drept conservatori). În privința susținerii candidatului de la alegerile prezidențiale din 2016, 29% au răspuns că l-au susținut pe Trump, 46% au spus că au susținut-o pe Clinton și restul au răspuns că nu au susținut niciunul dintre principalii doi candidați.

Design și procedură

Tratamentele (variabilele independente) constau în bazele de referință ale informației despre expunerea americanilor la fake news în timpul alegerilor prezidențiale din 2016. Estimările provin de la Guess et al. (2018). Participanții au fost repartizați aleatoriu într-unul din patru grupuri. Toate

grupurile au primit informația următoare: „Vă vom întreba care sunt presupunerile dumneavoastră despre prevalența fake news în perioada anterioară alegerilor prezidențiale din 2016 din SUA (7 octombrie — 14 noiembrie 2016)“. Primul grup ($n = 244$) a fost apoi expus la informații despre consum luând ca bază de referință numărul mediu de articole fake news pe care fiecare american le-a văzut în perioada anterioară alegerilor prezidențiale din 2016, formulate astfel: „Ca punct de referință, cercetările estimează că fiecare american cu drept de vot a văzut în medie cinci articole și jumătate de fake news în această perioadă“. Al doilea grup ($n = 251$) a văzut informații referitoare la consum având drept bază de referință procentul de americani expuși la orice articol de fake news în perioada anterioară alegerilor, formulate astfel: „Ca punct de referință, cercetările estimează că 27% dintre americanii cu drept de vot au accesat un articol de fake news în această perioadă“. Al treilea grup ($n = 232$) a văzut ambele formulări. Al patrulea grup ($n = 254$) a reprezentat grupul de control și nu a fost expus la nicio variabilă independentă.

Participanții au oferit date demografice și variabile moderatoare pre-tratament. După care, participanții au fost expuși la variabilele independente și ulterior au fost măsurate efectele percepției lor asupra extinderii și influenței fake news și atitudinilor lor față de fake news și de modalitățile de a le contracara.

Evaluari

Variabile dependente

Participanții au oferit o serie de estimări referitoare la expunerea la fake news și influența acestora, precum și la atitudinile față de fake news.

Consumul general perceput a fost evaluat prin întrebarea: „Cât de mult

credeți că au consumat americanii fake news în perioada alegerilor din 2016?“ cu răspunsuri pe o scală cu cinci puncte, de la „niciuna“ (1) la „multe“ (5), $M = 3,25$, $SD = 0,96$. Tendința percepută de consum (2016-prezent) a fost evaluată pe o scală cu trei puncte (1 = mai mică, 2 = a rămas aceeași, 3 = mai mare), $M = 2,49$, $DS = 0,64$.

Consumul perceput de grup a fost evaluat în cazul mai multor grupuri demografice distincte (vezi Grinberg et al., 2019; Guess et al., 2018; Guess et al., 2019). Cerința adresată participanților a fost: „Încercați să estimați câte știri false au consumat indivizii din următoarele grupuri“ în acea perioadă. Răspunsurile au fost notate pe o scală cu cinci puncte, de la „mult mai puține decât media“ la „mult mai multe decât media“. Grupurile erau susținătorii lui Trump ($M = 3,88$, $DS = 1,03$), susținătorii lui Clinton ($M = 3,52$, $DS = 1,02$), categoria de vârstă 18–34 ($M = 3,44$, $DS = 1,04$), categoria de vârstă 34–59 ($M = 3,02$, $DS = 0,91$), categoria de vârstă 60+ ($M = 3,62$, $DS = 1,17$), absolvenții de facultate ($M = 3,02$, $DS = 1,06$) și persoanele fără studii superioare ($M = 3,68$, $DS = 0,97$).

Participanții au oferit și „consecințe“ ce pot fi influențate de modificarea percepției asupra prevalenței știrilor false pe o scală Likert cu șapte puncte. Acestea au fost preocupările legate de efectele știrilor false ($M = 5,54$, $SD = 1,47$), convingerea că „știrile false sunt motivul pentru care Donald Trump a câștigat alegerile din 2016“ ($M = 4,17$, $DS = 2,07$), convingerea că „Facebook a luat măsuri semnificative de a limita știrile false de pe propria platformă“ ($M = 3,56$, $DS = 1,71$), „susținerea unei mai mari reglementări a știrilor false de către guvernul american“ ($M = 4,81$, $DS = 1,84$) și „acordul de investiție de bani publici în inițiative de educație în media digitale“ ($M = 4,90$, $DS = 1,68$). Acești itemi au format o scală cu un coeficient relativ redus de fidelitate ($\alpha = 0,63$), astfel că am examinat individual fiecare astfel de consecință.

Variabile moderatoare

Următoarele variabile au fost considerate ca având un potențial rol de

moderare a efectelor tratamentului informației referitoare la consum în funcție de baza de referință (procent din populație sau număr de articole).

Utilizarea de Facebook ($M = 6,02$, $SD = 2,57$) a fost evaluată pe o scală cu nouă puncte, de la „niciodată” la „aproape tot timpul”. Cunoștințele politice ($M = 2,80$, $SD = 1,14$) au fost evaluate prin scorul obținut pentru răspunsurile corecte la cinci întrebări: Câți ani rămâne în funcție un senator american?, Câți ani rămâne în funcție un membru al Camerei Reprezentanților?, Câți senatori provin din fiecare stat?, De câte ori un individ poate fi ales președinte? și Cine este în prezent prim-ministrul Marii Britanii? Interesul politic ($M = 3,35$, $SD = 1,08$) a fost măsurat pe o scală cu cinci puncte, de la „deloc interesat” la „extrem de interesat”.

Reflecția cognitivă a fost evaluată prin suma răspunsurilor corecte la trei itemi, folosind formulări alternative ale întrebărilor (Patel, 2017) pentru reducerea efectelor expunerii anterioare ($M = 1,61$, $DS = 1,20$). Necesitatea de a avea dovezi ($M = 5,77$, $SD = 1,06$) a fost reprezentată de media scorului la trei itemi pe o scală Likert cu șapte puncte, iar încrederea în intuiție ($M = 4,11$, $SD = 1,45$) a fost media a patru itemi pe o scală Likert cu șapte puncte (Garrett și Weeks, 2017).

Rezultate

Efecte asupra consumului perceput și interesului

Inițial am modelat efectele tratamentului asupra consumului general perceput, asupra tendinței percepute de consum (2016-prezent) și asupra fiecărei consecințe percepute individual, prin regresie OLS în care fiecare variabilă

independentă a fost introdusă ca variabilă binară, iar grupul de control a reprezentat categoria de referință. Aceste rezultate sunt prezentate în Tabelul 2.1.

Tabellul 2.1. Efectele tratamentului asupra consumului perceput și interesului

	Consum general			Tendință 2016-2018			Interes			Înfăscarea lui Trump			Acțiuni ale Facebook			Susținere a reglementării			Susținere a educației		
	b	SE	p	b	SE	p	b	SE	p	b	SE	p	b	SE	p	b	SE	p	b	SE	p
Baza de referință 5,5 articole	0,06	0,09	0,472	0,04	0,06	0,480	0,15	0,13	0,270	-0,14	0,19	0,456	-0,07	0,15	0,666	-0,18	0,17	0,288	-0,22	0,15	0,146
Baza de referință 27%	-0,03	0,09	0,743	0,08	0,06	0,162	0,19	0,13	0,138	-0,11	0,18	0,567	-0,03	0,15	0,838	-0,10	0,16	0,525	-0,34	0,15	0,024
Ambele baze de referință	-0,10	0,09	0,275	0,14	0,06	0,017	0,26	0,13	0,049	-0,08	0,19	0,684	0,00	0,16	0,984	-0,09	0,17	0,607	-0,09	0,15	0,024
Constantă	3,47	0,06	0,000	2,42	0,04	0,000	5,39	0,09	0,000	4,25	0,13	0,000	3,58	0,11	0,000	4,91	0,12	0,000	5,06	0,11	0,000
R ²	0,00			0,01			0,00			0,00			0,00			0,00			0,01		

Notă: N = 980

Această analiză nu arată niciun efect al bazei de referință asupra consumului general perceput (coloana 1). Însă expunerea simultană la ambele baze de referință (condiția 3) a determinat amplificarea percepției potrivit căreia consumul de știri false a crescut din 2016 (coloana 2), $b = 0,14$, $p = 0,017$. În mod similar, expunerea simultană la ambele baze de referință a determinat creșterea nivelului de interes pentru efectele știrilor false (coloana 3), $b = 0,26$, $p = 0,049$.

Rezultatele noastre nu arată niciun efect produs de bazele de referință asupra convingerii că știrile false au susținut victoria lui Trump, asupra convingerii că Facebook a luat măsuri semnificative pentru a limita răspândirea de știri false sau asupra susținerii reglementării guvernamentale a știrilor false (coloanele 4–6). Am descoperit că prezentarea procentului de americani expuși la știri false în 2016 a dus la reducerea suportului pentru programe de educație media finanțate public (coloana 7), $b = -0,34$, $p = 0,024$.

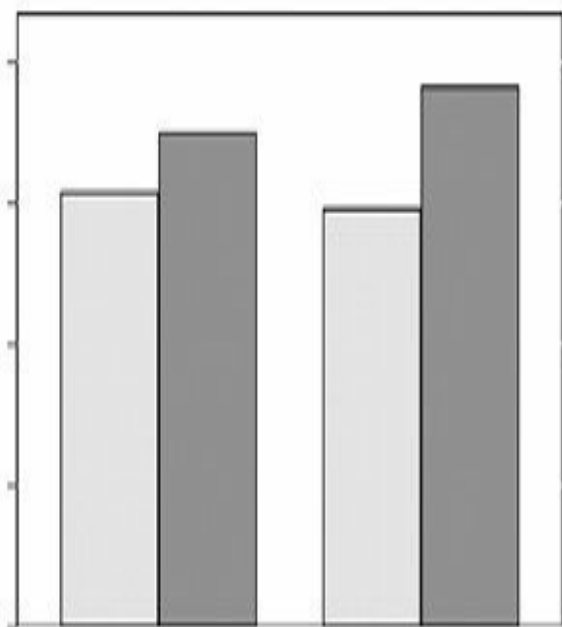
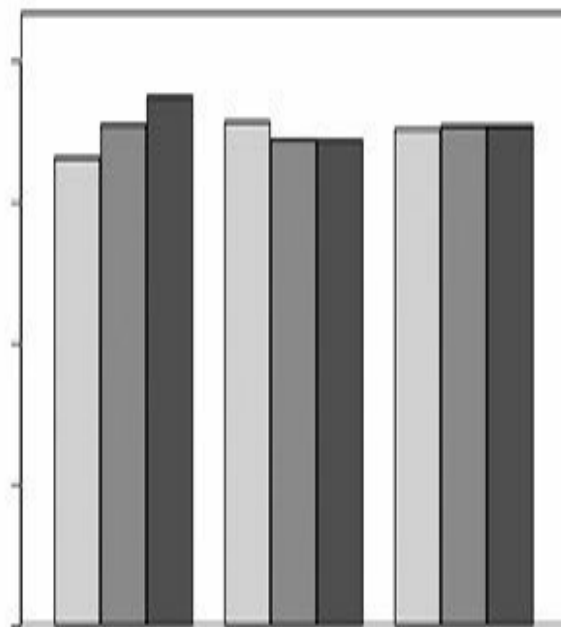
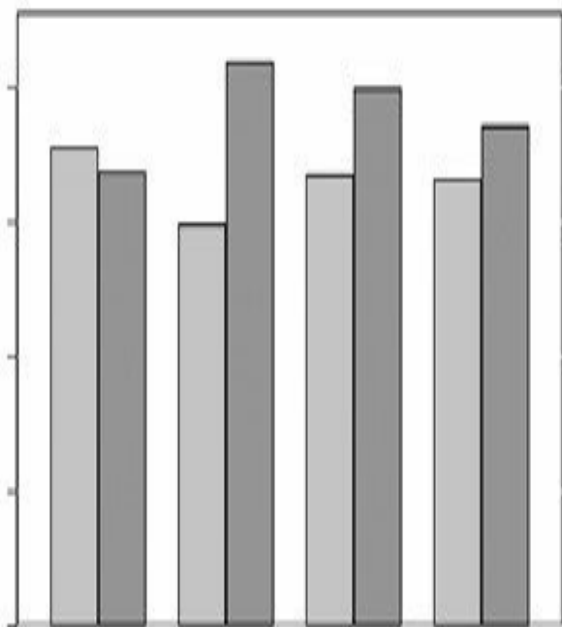
Examinând potențialii moderatori ai acestor efecte ale tratamentului, am descoperit că efectul expunerii simultane la ambele baze de referință asupra interesului din ce în ce mai mare pentru fenomenul fake news a fost semnificativ mai puternic în cazul susținătorilor lui Trump, $b = 0,83$, $p = 0,006$. Celelalte măsurători politice și cognitive nu au moderat efectele tratamentului.

Efecte asupra distorsionării grupo centrice

Am examinat apoi percepțiile referitoare la consum în cadrul diferitelor grupuri demografice. Prezintă-măi întâi în Figura 2.1 mediile pentru consumul perceput în interiorul grupului și în afara grupului. După cum se vede în această figură, susținătorii lui Clinton au estimat un consum

semnificativ mai mare de știri false în rândul susținătorilor lui Trump în comparație cu membrii propriului grup. În același timp, susținătorii lui Trump au estimat că susținătorii lui Clinton au consumat puțin mai multe știri false decât membrii propriului grup. Și absolvenții de facultate au estimat un consum semnificativ mai mare de știri false în rândul persoanelor fără studii superioare în comparație cu membrii propriului grup. Însă persoanele fără studii superioare au estimat la rândul lor un consum mai mare de știri false în cadrul propriului grup, deși într-o măsură mai mică. În sfârșit, participanții cu vârste între 18 și 24 de ani au estimat un consum mai mare de știri false în rândul americanilor cu vâsta de peste 60 de ani în comparație cu propriul grup, dar participanții de peste 60 de ani nu au prezentat percepția opusă (să observăm că în eșantionul nostru sunt semnificativ mai mulți tineri: 18–34, n = 513; 35–59, n = 422; 60+, n = 45).

Figura 2.1. Estimarea consumului de fake news în cadrul grupurilor demografice în cadrul grupului demografic al participantului



4

3

2

1

0

4

3

2

1

0

Donald Trump

Susținerea candidaților în 2016

Studii superioare

Facultate = 0

Hillary Clinton

grup participant

grup participant

grup participant

Alt candidat

Facultate = 1

Estimare consum facultate

Estimare consum Clinton

Niciun candidat

Estimare consum fără facultate

Estimare consum Trump

4

3

2

1

0

Vârsta

18–34

35–59

Estimare consum 18–34

60+

Estimare consum 35–59**Estimare consum 60+**

Am calculat apoi scorul diferențial al acestor percepții de grup (i.e. consumul perceput de susținătorii lui Trump — consumul perceput de susținătorii lui Clinton) în cadrul grupurilor demografice ale participanților, prezentate în Figura 2.2.

Figura 2.2. Diferența medie de estimare a consumului de fake news în grupurile demografice ale participanților

media (Trump-Clinton)

media (60+ — 18–34)

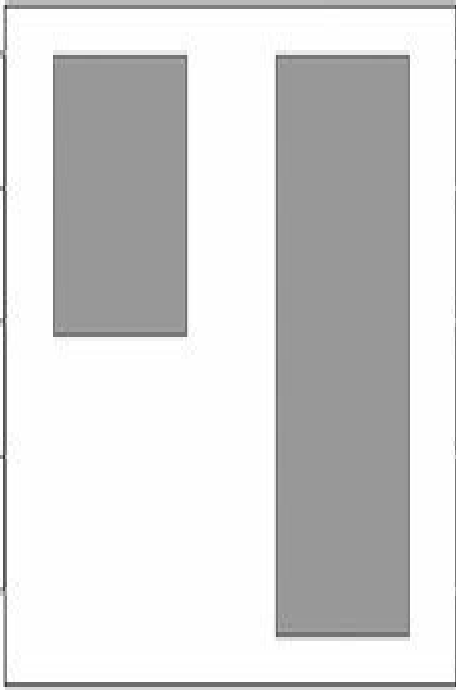
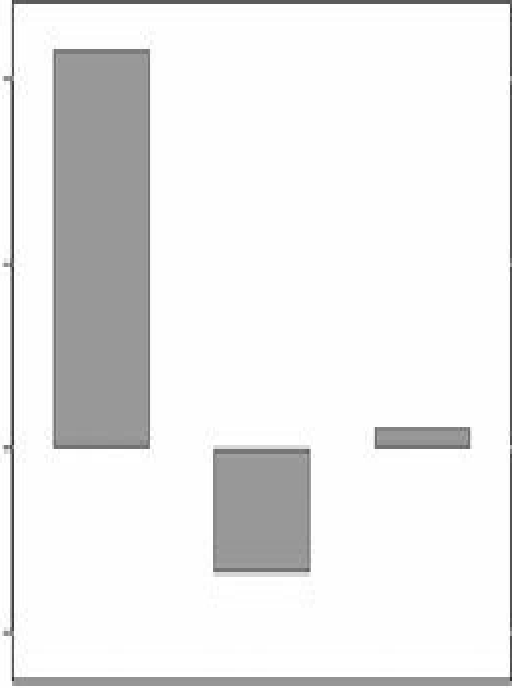
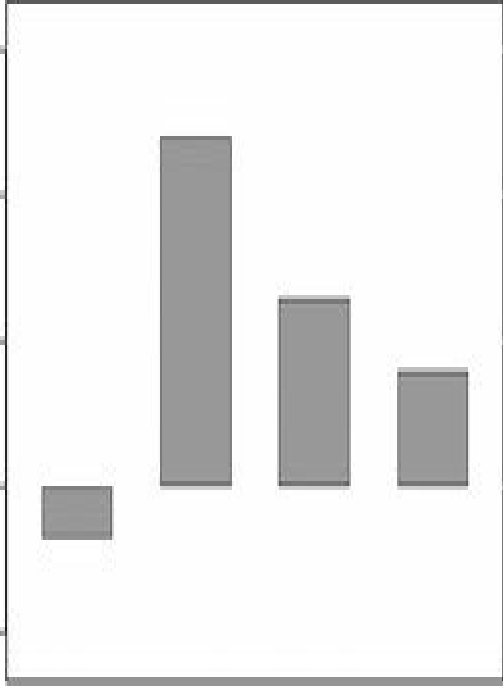
1,5

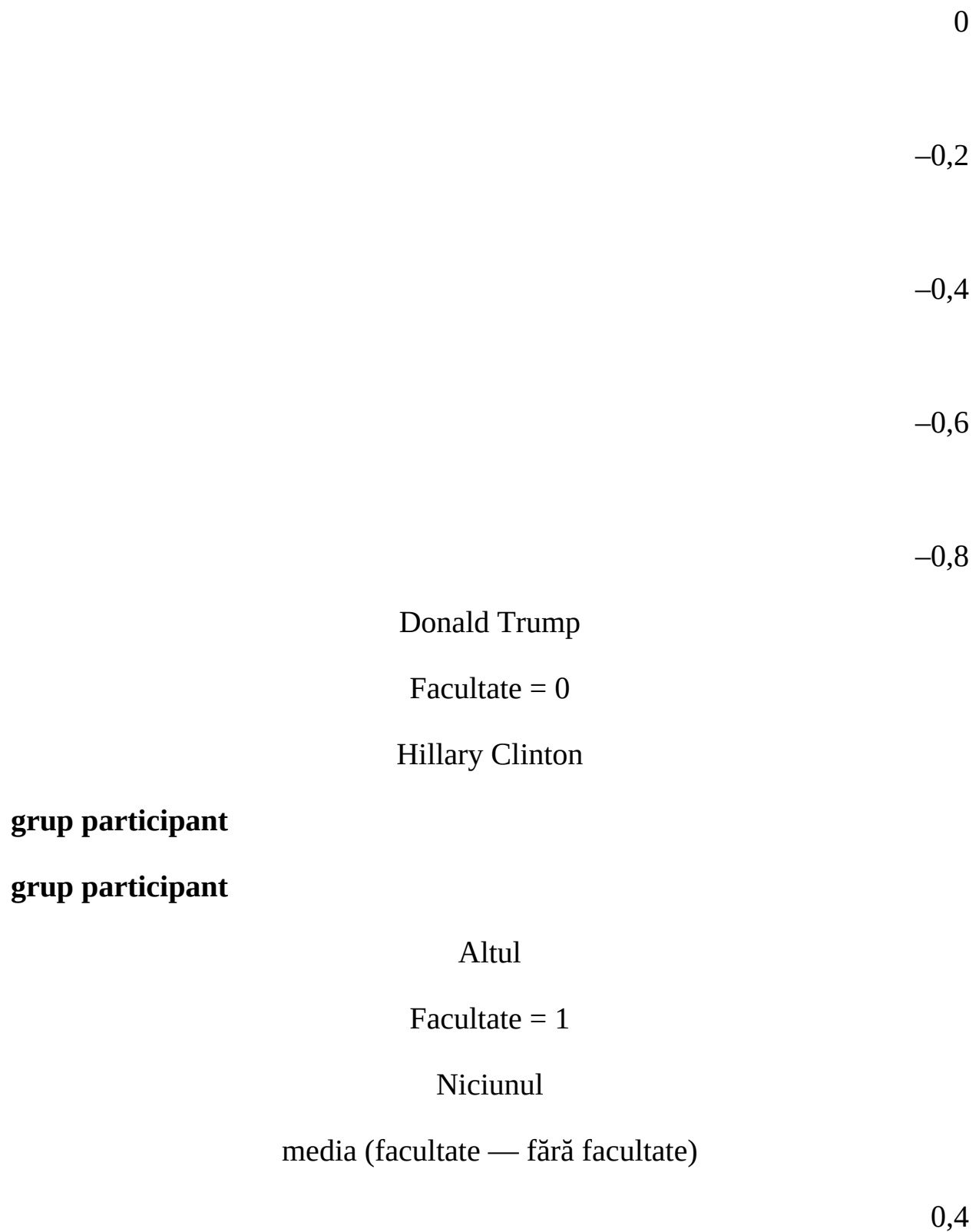
1

0,5

0

−0,5





	0,2
	0
	-0,2
18–34	
grup participant	
35–59	
60+	

Apoi am modelat consumul perceput de grupurile demografice utilizând regresia OLS. Am inclus ca predictorii tratamentele (variabilele noastre independente) și categoriile de apartenență la grup. Rezultatele sunt prezentate în Tabelul 2.2. În fiecare caz, apartenența la grup a prezis consumul perceput. Sprijinul pentru Trump (în comparație cu cei care nu susțineau niciun candidat) a prezis percepția unui consum mai redus de știri false de către susținătorii lui Trump ($b = -0,49$, $p < 0,001$) și percepția mai accentuată a consumului în rândul susținătorilor lui Clinton ($b = 0,21$, $p < 0,013$). Sprijinul pentru Clinton (în comparație cu cei care nu susțineau niciun candidat) a prezis percepția unui consum mai redus în rândul susținătorilor lui Clinton ($b = -0,35$, $p < 0,001$) și percepția unui consum mai mare în rândul susținătorilor lui Trump ($b = 0,32$, $p < 0,001$). Apartenența la categoria de vârstă 18–34 de ani (în comparație cu grupul 35–59 de ani) a prezis percepția unui consum mai mare de știri false în rândul celor de peste 60 de ani ($b = 0,31$, $p < 0,001$) și percepția unui consum mai redus în grupul cu vârste între 18 și 34 de ani ($b = -0,24$, $p < 0,001$). Absolvirea de studii superioare (în

comparație cu educația liceală sau gimnazială) a prezis percepția unui consum mai mare de către cei fără studii superioare ($b = 0,33$, $p < 0,001$) și percepția unui consum mai redus de către absolvenții de facultate ($b = -0,13$, $p = 0,064$).

Adunarea unei serii întregi de apartenențe ne arată că apartenența la grupurile nerelevante — în speță, susținerea unui candidat — prezice și percepția consumului de către celelalte grupuri avute în vedere. Susținerea lui Trump a prezis percepția unui consum mai mare de către categoria de vârstă 18–34 de ani ($b = 0,44$, $p < 0,001$) și percepția unui consum mai redus la grupa de vârstă de peste 60 de ani ($b = 0,44$, $p < 0,001$) după controlul vârstei, în timp ce predicția determinată de susținerea lui Clinton a avut sensul opus — un consum mai mare la categoriile de vârstă de peste 60 de ani ($b = 0,34$, $p < 0,001$) și 35–59 de ani ($b = 0,23$, $p = 0,002$). Și susținerea lui Trump a prezis percepția unui consum mai mare de către cei cu studii superioare ($b = 0,54$, $p < 0,001$), în timp ce susținerea lui Clinton a prezis percepția unui consum mai mare de către cei fără studii superioare ($b = 0,24$, $p = 0,002$) și un consum mai mic de către absolvenții de studii superioare ($b = -0,19$, $p < 0,019$) după controlul nivelului de educație. Astfel, percepția consumului de către diferitele grupuri demografice poate reflecta concepții diferite a ceea ce înseamnă „fake news“ în rândul susținătorilor lui Clinton și ai lui Trump.

Tabelul 2.2. Estimarea consumului grupurilor demografice în funcție de tratament și de caracteristicile demografice ale participanților

	Vârstă				
	60+	35–59	18–34		
	b	SE	p	b	SI
Baza de referință 5,5 articole	-0,11	0,10	0,31	0,04	0,
Baza de referință 27%	-0,11	0,10	0,28	-0,05	0,

Ambele baze de referință	−0,14	0,11	0,18	0,05	0,
18–34	0,31	0,08	0,00	0,11	0,
60+	0,11	0,18	0,54	0,11	0,
Constantă	3,54	0,08	0,00	3,44	0,
N	980				
R2	0,02				

Susținerea candidatului

	Clinton		Trump		
	b	SE	p	b	SE
Baza de referință 5,5 articole	0,01	0,09	0,91	−0,04	0,09

Am modelat apoi scoruri diferențiale utilizând aceeași procedură. Rezultatele sunt prezentate în Tabelul 2.3. Apartenența la grup a prezis și aceste scoruri diferențiale: sprijinul pentru Clinton ($b = 0,68$, $p < 0,001$) și sprijinul pentru Trump ($b = -0,71$, $p < 0,001$) au prezis fiecare scorul diferențial Trump-Clinton; categoria de vârstă 18–34 ($b = 0,56$, $p < 0,001$) a prezis scorul diferențial între cei tineri și ceilalți; iar absolvirea de studii superioare ($b = 0,45$, $p < 0,001$) a prezis scorul diferențial al nivelului de educație. În ce privește tratamentele, doar una dintre bazele de referință a prezis un scor diferențial: expunerea la numărul mediu de articole de știri false consumate de americani în 2016 a determinat reducerea diferenței percepute între americanii cu și fără studii superioare în privința consumului de știri false, $b = -0,37$, $p = 0,003$.

Tabelul 2.3. Scorurile diferențiale ale estimării consumului în funcție de tratament și de datele demografice ale participanților

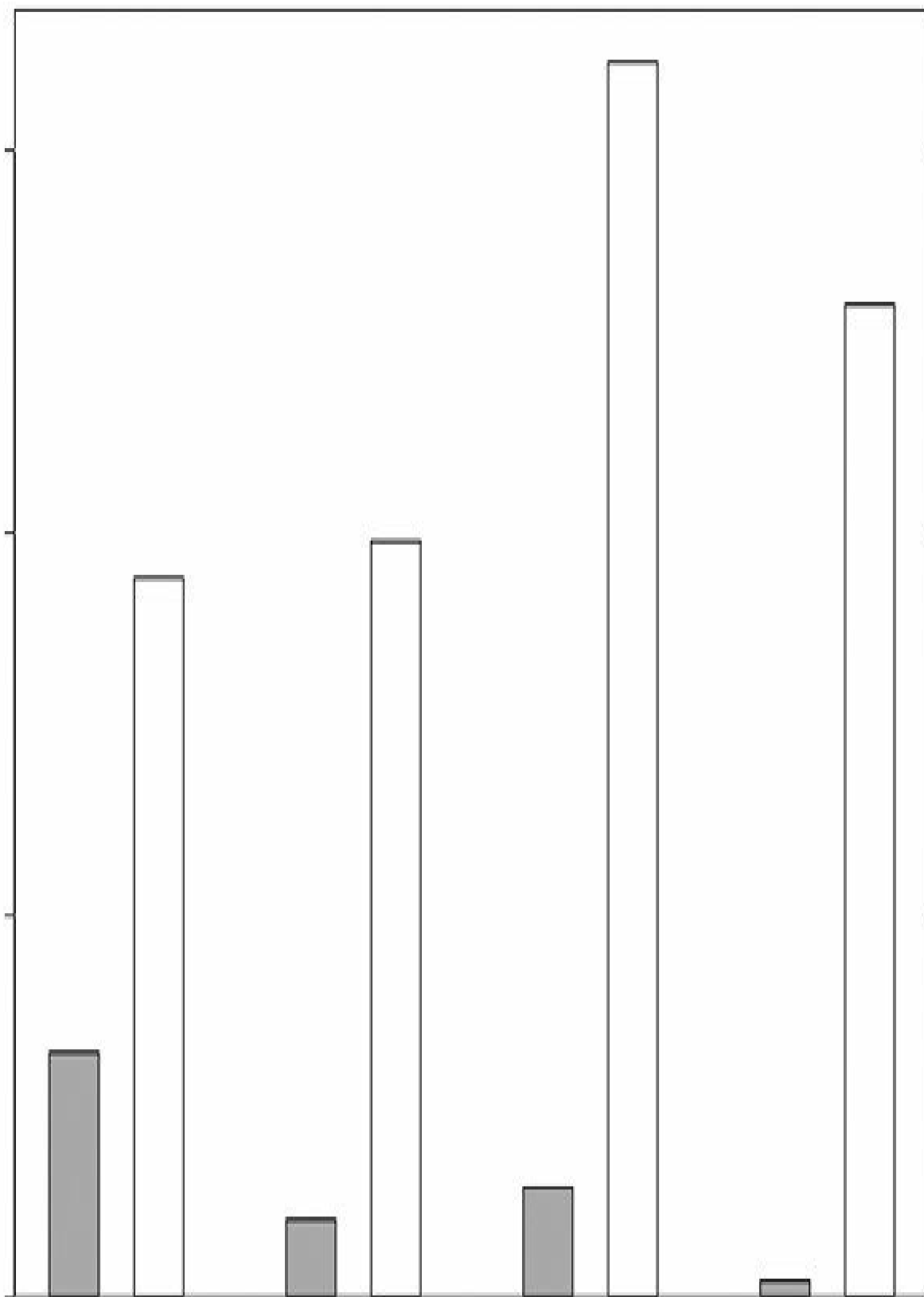
Apoi am inclus termenii interacțiunii dintre fiecare tratament al informației și fiecare categorie de apartenență la grup. Aceste rezultate sugerează că tratamentul informației a determinat creșterea scorului diferențial Trump-Clinton al susținătorilor lui Clinton, după cum se vede în Figura 2.3. În cazul participanților care au susținut-o pe Clinton în 2016, tratamentul ce a inclus informația referitoare la procentul de americani expuși la știri false ($b = 0,93$, $p = 0,004$) și tratamentul ce a inclus ambele baze de referință ($b = 0,77$, $p = 0,024$) au determinat creșterea diferenței dintre consumul perceput de știri false al susținătorilor lui Clinton și consumul perceput de știri false al susținătorilor lui Trump.

Discuție

Alarma publică provocată de știrile false poate fi disproporționată față de prevalența reală a expunerii la acest tip de conținut fabricat și a consumului acestuia. Cum poate faptul de a oferi statistici concrete ale consumului să influențeze evaluarea subiectivă a consumului de fake news și interesul pentru influența acestora? Servesc astfel de tratamente ale informației ca norme descriptive și duc la o „corectare“, reducând evaluările subiective? Sau determină amplificarea problemei fake news și accentuează aceste evaluări? Am descoperit că efectele bazei de referință a informației asupra percepției fake news, cel puțin în forma oferită de noi, sunt probabil reduse. Mai mult, fără o contextualizare adecvată, aceste informații pot mai degrabă să amplifice, și nu să reducă prevalența percepută, influența percepută și interesul pentru știri false. Expunerea simultană la ambele tipuri de bază de referință (procentul de americani expuși la știri false și media numărului de articole consumate) a amplificat percepția că știrile false au fost consumate mai mult după 2016 și a determinat creșterea interesului general pentru știri false. Nu am găsit dovezi semnificative că scorurile obținute la evaluările cunoștințelor de politică și ale abilităților cognitive ar modera efectele tratamentului.

Figura 2.3. Diferența mediilor consumului perceput de fake news în funcție de tratamente și de candidatul susținut

medie (consum al celor pro-Trump — consum al celor pro-Clinton)



1,5

1

0,5

0

Control

Metoda 5,5 articole

Metoda 27%

Ambele metode

Susținător non-Clinton

Susținător Clinton

Am găsit diferențe semnificative între grupurile considerate a consuma cele mai multe fake news. Deloc surprinzător, diferențele sunt determinate de apartenența indivizilor la diferite categorii sociale. Datele noastre prezintă distorsionări intragrupale accentuate în privința consumului perceput în categoriile preferinței la votul din 2016, a vârstei și a nivelului de educație al subgrupurilor (deși aceste diferențe de percepție nu sunt întotdeauna

simetrice).

Mai mult, apartenența non-relevantă la grup — specific, susținerea unuia dintre candidați — a prezis și consumul perceput al celorlaltor grupuri. Susținătorii lui Trump au presupus că persoanele tinere și educate consumau mai multe fake news, chiar și după ce au fost controlate vârsta și nivelul de educație, în timp ce susținătorii lui Clinton au presupus că persoanele în vârstă și cele fără studii superioare consumau mai multe fake news. Aceste relații indirecte pot reflecta faptul că publicul cunoaște bine bazele demografice ale sprijinului electoral pentru Clinton și Trump sau pot reflecta concepții diferite în privința a ceea ce constituie „fake news” (sau ambele variante). Se poate ca susținătorii lui Clinton să folosească o definiție mai apropiată de cea științifică a știrilor false (e.g. Guess et al., 2018), în timp ce susținătorii lui Trump, conform retoricii partidului, pot vedea publicații ca The New York Times și The Washington Post ca „fake news”. În unele cazuri (i.e. printre susținătorii lui Clinton), informațiile despre ratele consumului de știri false pot accentua diferența dintre consumul perceput în rândul susținătorilor lui Trump și Clinton.

Este important să observăm câteva limitări ale acestui studiu. În primul rând, dat fiind că eșantionul a fost recrutat folosind Amazon Mechanical Turk, care favorizează prezența persoanelor tinere, educate și liberale, suntem limitați în privința unor generalizări mai largi ale acestor descoperiri (dar vezi Coppock, 2019; Coppock, Leeper și Mullinix, 2018; Mullinix, Leeper, Druckman și Freese, 2015). Trebuie să ținem seama de acest aspect în special când avem în vedere rezultatele referitoare la diferențele de vârstă care reflectă diferențele observate între grupuri (neexpuse la bazele de referință). Cercetări viitoare ar trebui să reproducă aceste teste la nivelul unui eșantion de votanți reprezentativ la nivel național. Apoi, manipulările noastre au fost mai degrabă subtile. Nu am inclus o evaluare a acestora pentru a simula efectele unor astfel de baze statistice în medii de navigare media cu atenție limitată (e.g. Twitter). Se poate ca acest fapt să fi dus la efecte mai slabe asupra rezultatelor noastre de interes. De asemenea, în ce privește tratamentele noastre, este important să observăm că nu toate bazele de referință sunt egale (Westwood et al., 2019). Există probabil limite ale acestui tip de efect de proeminență ce teoretizăm că a avut loc prin aplicarea tratamentelor noastre. De exemplu, o bază de referință conform căreia un eveniment a influențat 5%

din populație poate avea o probabilitate mai mică de a determina creșterea interesului în comparație cu o bază de referință de 27%, așa că se poate ca euristica disponibilității să nu fie singurul mecanism al efectelor bazelor de referință, acestea depinzând în schimb de context. Studiile realizate în viitor ar trebui să evalueze efectele bazelor de referință pentru a determina dacă ancorarea joacă un rol în privința acestor efecte și ar trebui să testeze mecanismele într-un mod mai formal.

Cu toate astea, descoperirile noastre aduc informații noi referitoare la modul în care publicul gândește despre știrile false și le percepe. Estimările empirice derivate din cercetările realizate în științele sociale au o probabilitate redusă de a remodela într-un mod semnificativ percepțiile publicului asupra fake news ca problemă socială. Identitățile sociale ale publicului determină probabil stereotipul consumatorului de fake news și, sub acest aspect, ce anume este definit ca fake news. Așa stând lucrurile, inițiativele de educație în consumul media se pot concentra asupra comunicării bazelor a ceea ce constituie jurnalism legitim.

Fake news sunt considerate de mulți ca reprezentând un pericol concret și actual la adresa democrației. Uniunea Europeană a organizat un grup la nivel înalt pentru a studia această problemă și mai multe guverne naționale au organizat audieri sau au format comitete pentru examinarea acesteia. Rezultatele prezentate în acest capitol sugerează că schimbarea convingerilor despre consumul de fake news ar putea fi dificilă. Această descoperire ar putea reprezenta o problemă.

Dat fiind că suntem cercetători în domeniul dezinformării, cunoaștem prea bine efectele distructive pe care informațiile eronate le pot avea la nivelul politicii și societății. Cu toate astea, efectele negative ale fake news trebuie evaluate atent. Deși informațiile eronate sunt într-adevăr o problemă serioasă la adresa societății, este important să nu fim distrași de aceste forme mai degrabă de nișă și să lăsăm informațiile eronate să se răspândească sub alte forme. De exemplu, informațiile false comunicate de elite — în special când au acces direct la audiențe mult mai mari prin social media — au o probabilitate mai mare de a constitui o problemă serioasă în comparație cu site-urile de fake news. Ar fi păcat dacă eforturile noastre de a restrânge informațiile false ar rezolva problema fake news, dar ar lăsa astfel cale liberă

altor forme de inducere în eroare și disimulare.

Referințe

Ahler, D.J. și Sood, G. (2018). The parties in our heads: Misperceptions about party composition and their consequences. *The Journal of Politics*, 80 (3), 964–981.

An, S. (2008). Antidepressant direct-to-consumer advertising and social perception of the prevalence of depression: Application of the availability heuristic. *Health Communication*, 23 (6), 499–505.

Barthel, M., Mitchell, A. și Holcomb, J. (2016). Many Americans believe fake news is sowing confusion. Pew Research. Accesat la www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/

Carey, B. (2018, January 2). „Fake news“: Wide reach but little impact, study suggests. *The New York Times*. Accesat la www.nytimes.com/2018/01/02/health/fakenews-conservative-liberal.html

Carroll, J.S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of Experimental Social Psychology*, 14 (1), 88–96.

Coppock, A. (2019). Generalizing from survey experiments conducted on MechanicalTurk: A replication approach. *Political Science Research and Methods*, 7 (3), 613–628.

Coppock, A., Leeper, T.J. și Mullinix, K.J. (2018). Generalizability of heterogeneous treatment effect estimates across samples. *Proceedings of the National Academy of Sciences*, 115 (49), 12441–12446.

Ecker, U.K., Lewandowsky, S., Chang, E.P. și Pillai, R. (2014). The effects of subtle misinformation in news headlines. *Journal of Experimental Psychology: Applied*, 20 (4), 323.

Folkes, V.S. (1988). The availability heuristic and perceived risk. *Journal of Consumer Research*, 15 (1), 13–23.

Fox, A. (ianuarie 2019). Majority of Americans were not exposed to „fake news“ in 2016 U.S. election, Twitter study suggests. Accesat la www.sciencemag.org/news/2019/01/majority-americans-were-not-exposed-fake-news-2016-us-electiontwitter-study-suggests

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19 (4), 25–42.

Garcia-Retamero, R. și Galesic, M. (2009). Communicating treatment risk reduction to people with low numeracy skills: A cross-cultural comparison. *American Journal of Public Health*, 99 (12), 2196–2202.

Garrett, R.K. și Weeks, B.E. (2017). Epistemic beliefs' role in promoting misperceptions and conspiracist ideation. *PloS One*, 12 (9), e0184733.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B. și Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science*, 363 (6425), 374–378.

Guess, A., Lyons, B.A., Montgomery, J., Nyhan, B. și Reifler, J. (2019). Fake news, Facebook ads, and misperceptions: Assessing information quality in the 2018 U.S. midterm election campaign. Accesat la www.dartmouth.edu/~nyhan/fakenews-2018.pdf

Guess, A., Nagler, J. și Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, 5 (1), eaau4586.

Guess, A., Nyhan, B. și Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. Accesat la

www.dartmouth.edu/~nyhan/fake-news-2016.pdf

Hill, S.J. (2017). Learning together slowly: Bayesian learning about political facts. *The Journal of Politics*, 79 (4), 1403–1418.

Hoffrage, U., Lindsey, S., Hertwig, R. și Gigerenzer, G. (2000). Communicating statistical information. *Science*, 290, 2261–2262. doi: 10.1126/science.290.5500.2261

Kessler, S., Levine, E.K., Opitz, J.M. și Reynolds, J.F. (1987). Psychological aspects of genetic counseling: IV. The subjective assessment of probability. *American Journal of Medical Genetics*, 28 (2), 361–370.

MacLeod, C. și Campbell, L. (1992). Memory accessibility and probability judgments: An experimental evaluation of the availability heuristic. *Journal of Personality and Social Psychology*, 63 (6), 890.

Mullinix, K.J., Leeper, T.J., Druckman, J.N. și Freese, J. (2015). The generalizability of survey experiments. *Journal of Experimental Political Science*, 2 (2), 109–138.

Neighbors, C., Dillard, A.J., Lewis, M.A., Bergstrom, R.L. și Neil, T.A. (2006). Normative misperceptions and temporal precedence of perceived norms and drinking. *Journal of Studies on Alcohol*, 67 (2), 290–299.

Neisser, U. (1978). Memory: What are the important questions? În M.M. Gruneberg, P.E. Morris și H.N. Sykes (coord.), *Practical aspects of memory* (pp. 3–24). Londra: Academic Press. [Retipărit în Neisser, U. (1982). *Memory: What are the important questions. Memory Observed: Remembering in Natural Contexts*, 3–19].

Patel, N. (2017). The cognitive reflection test: A measure of intuition/reflection, numeracy, and insight problem solving, and the implications for understanding real-world judgments and beliefs. (Teză de doctorat nepublicată). Columbia, MO: University of Missouri.

Rimal, R.N. și Real, K. (2005). How behaviors are influenced by perceived norms: A test of the theory of normative social behavior. *Communication*

Research, 32 (3), 389–414.

Schapira, M.M., Davids, S.L., McAuliffe, T.L. și Nattinger, A.B. (2004). Agreement between scales in the measurement of breast cancer risk perceptions. *Risk Analysis: An International Journal*, 24 (3), 665–673.

Turner, J.C., Brown, R.J. și Tajfel, H. (1979). Social comparison and group interest in ingroup favouritism. *European Journal of Social Psychology*, 9 (2), 187–204.

Tversky, A. și Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185 (4157), 1124–1131.

Westwood, S., Messing, S. și Lelkes, Y. (2019). Projecting confidence: How the probabilistic horse race confuses and demobilizes the public. Accesat la https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3117054.

Capitolul 3. Adevărul și dinamica difuzării de știri pe Twitter

Robert Ackland și Karl Gwynn

Introducere

Acest capitol examinează două aspecte ale informației false: cum se poate stabili dacă informația (de exemplu, o știre) este adevărată și cum adevărul sau neadevrul informației influențează difuzarea sau răspândirea acesteia. Capitolul se concentrează în special asupra semnificației pe care o are social media în privința informațiilor false (în special în ce privește fake news): prevalența și efectul acestora, precum și metodele de identificare și studiere ale fenomenului. Analizăm literatura recentă referitoare la modul în care metodele computaționale și sursele „big data” (e.g. social media) sunt utilizate pentru identificarea informațiilor false și pentru înțelegerea modului în care oamenii se lasă prinși de informațiile false și contribuie la răspândirea lor.

Aplicația noastră empirică implică o nouă abordare a verificării manuale a adevărului știrilor și aplicăm această metodă asupra unui eșantion de știri politice australiene din 2017. Explorăm apoi modul în care veridicitatea știrilor influențează difuzarea acestora pe Twitter, concentrându-ne asupra indicatorilor esențiali ai difuzării: răspândire (câți oameni sunt implicați în difuzare), viteză și extindere (cât de departe ajung știrile în rețea și cât de diferiți sunt actorii implicați în difuzare).

Fundal

În această secțiune analizăm mai întâi definițiile existente ale informației false (misinformation) și ale fake news și rezumăm apoi motivele pentru care oamenii contribuie la răspândirea informațiilor false. Trecem apoi în revistă cercetările asupra modului în care social media au potențialul de a amplifica problema fake news. Abordările centrate pe verificarea veridicității știrilor sunt apoi expuse și analizate sub aspectul capacității lor de a oferi evaluări acurate ale veridicității, de interes deosebit fiind abordările manuale (verificări realizate de experți în domeniu) și abordările computaționale. Trecem în revistă și unele dintre cercetările esențiale care folosesc metode computaționale și big data pentru a înțelege fenomenul informațiilor false.

Definirea informațiilor false și a fake news

Informațiile false (misinformation) sunt informațiile neadevărate care au capacitatea de a se răspândi la nivelul societății și de a influența opinia publică. Exemple sunt știrile satirice (create pentru umor, și nu cu scopul de a induce în eroare), zvonurile (informații care nu au fost confirmate încă a fi adevărate sau false), teorii ale conspirației (care prin definiție nu sunt verificabile și au tendința de a fi răspândite de persoane care le cred adevărate) și farse (create pentru a induce în eroare, pot fi amuzante sau răuvoitoare și adesea implică citarea unei surse de încredere). Informația falsă, la fel ca toate tipurile de informații, este graduală și trebuie revizuită când devin disponibile noi cunoștințe.

Utilizarea din prezent a termenului misinformation a fost politizată în urma

unor incidente precum „certificatul de naștere al lui Obama“ și „rețeaua de prostituție a minorilor a lui Clinton“. Deși misinformation se poate referi la orice informație eronată accesibilă publicului, utilizarea contemporană a termenului implică în general reavoință și inducere în eroare.¹ Pentru scopurile urmărite în acest capitol, misinformation se referă la toate cazurile în care informația poate fi dovedită ca incluzând în mod evident neadevăruri.

Fake news reprezintă un tip de informație falsă (misinformation) în care informația se referă la un eveniment care apare la știri, iar reavoința aparține persoanei (sau persoanelor) care creează știrile, dar nu neapărat și celor care distribuie știrile. Allcott și Gentzkow (2017) definesc fake news ca articole de știri care sunt „intenționat și verificabil false“. Producerea de fake news este în general motivată de câștiguri financiare sau ideologice, dat fiind că induc cititorii în eroare în privința calității de adevăr a conținutului știrilor false. Cu toate astea, termenul fake news este uneori utilizat pentru a desemna publicații legitime ca metodă de discreditare și defăimare (aceasta este o tactică aplicată adesea de președintele SUA Donald Trump), ceea ce indică volatilitatea termenului și susceptibilitatea acestuia de a fi folosit greșit.²

Motivații de răspândire a informațiilor false

De ce răspândesc oamenii informații lipsite de acuratețe? Adesea, ei cred că informația este adevărată din cauza distorsiunilor euristice. Erorile euristice desemnează tendințele indivizilor de a se baza pe tipare simpliste pentru a reduce consumul pretins de gândirea critică. Acest lucru este evident în cazul încrederii în credințe și opinii: dacă informația le confirmă, este mai probabil să fie crezută (distorsiune de confirmare) și, astfel, distribuită din nou (e.g. Ecker, Lewandowsky, Fenton și Martin, 2014). Lewandowsky, Ecker, Seifert, Schwarz și Cook (2012) au identificat patru factori care influențează faptul dacă o persoană crede informația: consecvența mesajului (este consecvent cu credințele anterioare?), coerența mesajului (are coerență

internă și e plauzibil?), credibilitatea sursei și acceptabilitatea generală (câți alți oameni par să o creadă?). Un alt motiv de răspândire a informațiilor false se referă la presiunile normative care îi fac pe oameni să răspândească informații false pentru a câștiga statut social și acceptare, conform teoriei identității sociale (e.g. Tajfel și Turner, 2001, 2004). Odată ce o persoană crede informația falsă, convingerile sunt dificil de schimbat (Lewandowsky et al., 2012), iar încercările de corectare a falsurilor pot chiar să amplifice răspândirea informațiilor false, în special în cadrul grupurilor ideologice (Nyhan și Reifler, 2010). Aceste lucruri confirmă necesitatea detectării acurate și din timp a informației false și au motivat construirea de sisteme de detectare a acesteia.

Fake news în era digitală

Deși fake news nu reprezintă un fenomen nou, există mai multe motive pentru importanța și pentru interesul din ce în ce mai mare care li se acordă în era digitală (vezi e.g. Allcott și Gentzkow, 2017; Shu, Sliva, Wang, Tang și Liu, 2017).

În primul rând, barierele în calea accesului pe piața știrilor au devenit semnificativ mai reduse, dat fiind că website-urile pot fi construite cu ușurință și pot fi transformate în surse de venit prin publicitate. În privința procesului de răspândire a fake news, costurile fixe asociate accesului la social media sunt foarte mici — ceea ce amplifică viabilitatea strategiilor pe termen scurt de constituire a prezenței pe social media a unei anume campanii de fake news și reduce motivația de construire a unei prezențe pe termen lung asociată jurnalismului serios. În al doilea rând, social media sunt adecvate diseminării de fake news; formatul social media e de așa natură, încât informația tinde să fie distribuită prin fragmente scurte de text, ceea ce face ca utilizatorii să găsească dificilă evaluarea veridicității. În al treilea rând, are loc o reducere constantă a încrederii publice și a încrederii în mass-media. În al patrulea rând, în multe state occidentale s-a amplificat polarizarea politică

(gradul de sentimente negative față de capătul opus al spectrului politic), ceea ce poate determina creșterea probabilității ca fake news să fie crezute.

Amplificarea polarizării politice este legată de un alt aspect important al social media care poate influența gradul în care oamenii sunt expuși la fake news. În primele cercetări asupra comunicării politice, Katz și Lazarsfeld (1955) au susținut că indivizii cu bune cunoștințe de media („liderii de opinie”) sunt intermediari între mass-media și public (teoria „comunicării în doi pași”). Însă social media au determinat o reducere a prezenței sau importanței intermediarilor dintre producători și consumatori de informație (concepție denumită „dezintermediere”), existând acum posibilitatea ca oamenii să creeze și să distribuie informații pe rețelele de socializare online. Este mult mai simplă acum selectarea fiecărei surse de informații (e.g. urmărirea pe Twitter a utilizatorilor care au aceleași convingeri politice), iar această „îngustare” poate duce la crearea de așa-numite camere de ecou: grupuri de utilizatori care au aceleași păreri și care nu sunt expuși la perspective diferite, ceea ce poate duce la o polarizare și mai mare (diferențe de atitudini). Legat de acestea este fenomenul „bulelor filtrante”: algoritmi utilizați de companiile de social media pentru a selecta conținutul nou pentru utilizatori în funcție de interacțiunile lor anterioare cu alte conținuturi, întărind astfel tiparele de consum de informație și reducând probabilitatea ca utilizatorii să fie expuși la informații noi.³

Camerele de ecou și bulele filtrante sunt relevante pentru problema fake news pentru că pot influența probabilitatea ca o persoană să transmită știri false în următoarele moduri. În primul rând, există o probabilitate mai mare ca un individ să fie conectat cu alte persoane care cred evident aceleași știri (ceea ce susține credibilitatea și întărește convingerile). În al doilea rând, individul are o probabilitate mai mare de a fi expus la știri false și s-a descoperit că o expunere mai mare determină o probabilitate mai mare de a crede (Hasher, Goldstein și Toppino, 1997) și o probabilitate mai mică de a fi expus la informații care ar contrazice o știre falsă aliniată ideologic. În sfârșit, există o presiune normativă din ce în ce mai mare de a răspândi fake news (chiar dacă persoana nu le crede).

Allcott și Gentzkow (2017) au descoperit că știrile false au înclinat puternic în favoarea lui Donald Trump în 2016; baza lor de date conține 115 știri false

pro-Trump care au fost distribuite pe Facebook de 30 de milioane de ori și 41 de știri false pro-Clinton care au fost distribuite pe Facebook de 7,6 milioane de ori. Guess, Nyhan și Reifler (2018) au studiat modul în care identitatea politică a influențat consumul de fake news în timpul alegerilor prezidențiale din 2016 din SUA. Autorii au descoperit că susținătorii lui Trump aveau o tendință disproporționată de a vizita website-uri care publicau fake news, și asta din cauza faptului că știrile false erau direcționate către susținătorii lui Trump, fiind așadar consecvente atitudinilor acestora (având astfel o probabilitate mai mare de a fi consumate).⁴

Pericolele pe care le comportă fake news constau în faptul că pot reduce încrederea în instituțiile democratice, pot reduce coeziunea socială și pot contribui la succesul liderilor populiști (unii comentatori au atribuit un rol semnificativ știrilor false în alegerea lui Donald Trump în 2016). Cu toate acestea, Vargo, Guo și Amazeen (2018) văd un alt efect potențial al fake news — capacitatea de a modela peisajul știrilor online. Autorii au utilizat cadrul conceptual oferit de Network Agenda-Setting (NAS), care susține că media de știri pot influența modul în care membrii publicului fac legătura între subiecte (e.g. criza energetică este pusă în legătură cu problema relațiilor externe) și pot scoate în evidență sau pot populariza anumite subiecte. Fake news pot avea un efect asupra stabilirii agendei publicațiilor de știri (în special a celor partizane) pur și simplu prin generarea de informații false (misinformation) cărora jurnaliștii trebuie să le răspundă (e.g. prin verificarea faptelor). Cercetătorii au evaluat potențialul efect al știrilor false asupra stabilirii agendei utilizând GDELT Global Knowledge Graph, o rețea de oameni, locuri, teme și evenimente construite computațional pe baza știrilor globale (Leetaru și Schrod, 2013). Folosind testele de cauzalitate ale lui Granger⁵, autorii au descoperit că știrile false au reușit să determine ieșirea în evidență în media online a anumitor subiecte (e.g. relații internaționale) în perioada 2014–2016. De asemenea, cercetarea a izolat publicațiile media partizane și a descoperit că publicațiile conservatoare au scos în evidență publicațiile de fake news, ceea ce a avut un efect asupra stabilirii agendei publicațiilor liberale (care răspundeau știrilor false).

Perspectivă de ansamblu asupra abordărilor verificării știrilor

Există două mari abordări pentru evaluarea veridicității știrilor.⁶ Prima este centrată pe conținutul în sine — aspecte ale textului (și ale imaginilor, dacă există) știrilor, precum și sursele sale. Această abordare de evaluare a veridicității știrilor în funcție de conținut poate fi subîmpărțită în abordări centrate pe cunoaștere și abordări centrate pe stil.

Abordările centrate pe cunoaștere implică utilizarea de surse externe pentru verificarea afirmațiilor pe care le fac știrile („fact checking“) și există trei variante de verificare a faptelor.

(1) Verificarea faptelor de către experți în domeniu (capitolul de față implică o implementare a verificării realizate de experți, așadar oferim o trecere în revistă a literaturii relevante în următoarea secțiune).

(2) Crowdsourcing este verificarea faptelor prin consensul agregat al membrilor publicului general. Aplicații ca Fiskkit și contul LINE „For real“ permit ca sugestii și comentarii ale utilizatorilor să ofere o indicație a veridicității știrilor și există un interes considerabil pentru utilizarea de registre partajate (distributed ledger technology — DLT) de tip blockchain pentru verificarea de fapte prin crowdsourcing.

(3) Bazele de cunoștințe (knowledge graphs) ca DBpedia sau Google Knowledge Graph sunt rețele care prezintă legăturile dintre entități din lumea reală (oameni, locuri, lucruri). Verificarea faptelor poate fi realizată automat prin compararea conținutului știrilor (și, în special, a relației dintre entitățile menționate în știri) cu conținutul existent în knowledge graph. Principala problemă a acestei metode este că, deși e rapidă și acurată, se bazează pe surse de informații care nu pot încorpora întreaga cunoaștere.

Abordările de verificare a veridicității știrilor în funcție de stil implică analiza lingvistică a textului știrilor, ceea ce include aspecte lexicale (e.g. număr de cuvinte, lungimea medie a cuvintelor, numărul de cuvinte unice) și aspecte sintactice (e.g. n-gram****, părți de vorbire). Pot fi utile și aspecte lingvistice specifice domeniului, cum ar fi link-uri externe și prezența de tabele și grafice. Imaginile pot, de asemenea, să fie utilizate pentru evocarea anumitor reacții afective (furie, șoc), care cresc probabilitatea de a crede știrea, astfel încât și aspectele extrase pe baza imaginilor sunt folosite în detectarea de știri false.

A doua abordare de verificare a veridicității știrilor presupune utilizarea de date despre contextul social al știrilor, adică despre modul în care sunt consumate și distribuite. Acesta este punctul în care social media au avut un efect major, dat fiind că permit analiza detaliată a contextului social în care știrile sunt consumate și răspândite. În primul rând, aspecte legate de utilizatorii care au interacționat cu știrile pot fi utilizate pentru detectarea fake news. În cazul Twitter, astfel de aspecte pot include numărul de prieteni/followers, vechimea contului utilizatorului și numărul de tweet-uri scrise. În al doilea rând, în social media este posibilă și colectarea de date post-level — reacțiile pe care le-au avut oamenii la articolele de știri (e.g. tweet-uri citate odată cu redistribuirea de link-uri de știri, comentarii la postări pe Facebook) — iar acestea pot oferi informații utile pentru detectarea fake news. Aceste reacții pot fi analizate în funcție de aspectele lingvistice, după cum am menționat mai sus, dar există aici și posibilitatea de utilizare a datelor social media în dezbaterile sau contestarea unei anumite știri. În al treilea rând, există aspecte ale rețelei care pot fi utilizate pentru detectarea știrilor false, de exemplu, rețele de difuzare a știrilor pe Twitter prin retweet-uri (retweet-uri în cascadă).

Verificarea faptelor de către experți în domeniu

Verificarea faptelor este utilizată adesea ca metodă de determinare a veridicității afirmațiilor și știrilor, în special dacă există un risc perceput de apariție a informațiilor false. Verificarea faptelor a ajuns în prim-plan în ultimii zece ani pe măsură ce website-uri ca Politifact.com și Factcheck.org s-au consacrat ca verificatori politici. Includerea recentă a verificării faptelor în timpul relatării în media a alegerilor prezidențiale din SUA din 2016 și pe durata președinției lui Trump subliniază capacitatea verificării faptelor de a se împotrivi informațiilor și știrilor false. Sistemul politic australian are propriile publicații de verificare a faptelor, cum ar fi RMIT University-ABC Fact Check și The Conversation's FactCheck.

Industria verificării faptelor s-a extins de la o atracție minoră în timpul ciclurilor electorale la un element proeminent al sferei politice. Young, Jamieson, Pulsen și Goldring (2018) au identificat alegerile prezidențiale din 1988 ca primul exemplu de aplicare contemporană a verificării faptelor în scopul monitorizării comportamentului potențialilor candidați la președinție. Așa a apărut verificarea veridicității informațiilor din campaniile electorale (ad watch), televiziunile transmițând cunoștințe provenite de la oamenii de știință și de la profesioniștii din media. Ceea ce a început ca notă de subsol a căpătat forță pe măsură ce mediul politic american a adoptat tehnici retorice menite a induce în eroare și a oferi informații false publicului.

Răspândirea informațiilor false din domeniul politic a constituit o zonă semnificativă de dezbateră științifică în anii 2000, în special în privința invaziei SUA în Irak și a presupusei dețineri de arme de distrugere în masă. Descoperirea campaniei de dezinformare coordonată de guvern a stimulat cercetările asupra modurilor în care indivizii deosebesc adevărul de fals și asupra ușurinței cu care politicienii înșală publicul.

Nyhan și Reifler (2010) fac o distincție binară între subiectul „neinformat” și subiectul „dezinformat”: subiecții neinformați reacționează prin distorsionări euristice, în timp ce subiecții dezinformați au primit informații false prin afiliere politică. Anularea dezinformării prin „informații corective” în general nu reușește, dat fiind că subiecții au încorporat cunoștințele în identitatea lor politică. Aceasta este o provocare mult mai mare decât un public neinformat, a cărui lipsă de cunoaștere nu este legată de concepte de identitate politică și

socială.

Verificarea faptelor a apărut simultan cu dezvoltarea rapidă a internetului și a ajuns o prezență din ce în ce mai mare online, întrucât putea oferi rapid reacții la informațiile politice false. Factcheck.org a fost lansat în 2003 de Annenberg Public Policy Center, cu scopul de a contrazice informațiile false prin capacitatea combinată a jurnaliștilor și a oamenilor de știință. Succesul pe care l-a avut Factcheck.org a dus la crearea și a altor website-uri, ca Politifact și Fact Checker (Washington Post), cu metode similare de detectare a informațiilor false.

Până în acest punct, verificarea faptelor a rămas o sursă de informații în general bine văzută, care a determinat creșterea responsabilității politice. Deși eficiența verificării faptelor în ceea ce privește implicarea audienței a fost desconsiderată de autori ca Nyhan și Reifler (2010), fundamentele epistemologice și metodologice nu au fost explorate extensiv. Uscinski și Butler (2013) oferă o critică a practicilor de verificare a faptelor și susțin că: „Aceste practici pornesc de la presupunerea tacită că nu poate exista dezbatere politică autentică despre fapte“ (p. 163).

Uscinski și Butler (2013) concluzionează că verificarea faptelor implică o interacțiune complexă între politică, societate, economie și istorie, reducând-o la măsurătorile cele mai simpliste, care stabilește adevărul într-un spectru unidimensional. Dacă se adoptă o epistemologie radical interpretativă, adevărul nu poate fi dictat de politicieni sau de cei care verifică faptele, într-o realitate complexă, care permite concepte diverse ale adevărului. Autorii subliniază neajunsuri metodologice foarte importante inerente verificării faptelor, definite de o epistemologie obiectivistă simplistă. Acestea includ efecte de selecție (selectarea faptelor se bazează de obicei pe subiecte senzaționaliste sau de pe prima pagină), aspecte multiple ale faptelor sau ale unei părți a faptelor (definind parametrii măsurabili ai unui singur fapt), simple afirmații (presupuneri în privința unei relații necunoscute între fapte), predicții și criterii de selecție (Ce anume constituie adevărul?, Cât de adevărat trebuie să fie ceva anume?). Autorii se abțin de la oferirea unui răspuns simplu la neajunsurile metodologiei verificării faptelor. În schimb, ei subliniază din nou natura interpretativă a faptelor și natura complexă a retoricii și informațiilor politice.

Amazeen (2015) adoptă o abordare critică în privința generalizărilor și alegerilor folosite de Uscinski și Butler (2013) pentru a evidenția neajunsurile verificării faptelor. Deși recunoaște natura interpretativă a verificării faptelor, Amazeen (2015) sugerează că Uscinski și Butler (2013) au ales exemple de verificări ale faptelor care susțin în mod evident opinia lor. Amazeen (2015) afirmă că multe fapte nu se supun dezbaterii și că faptele interpretabile constituie, în realitate, o minoritate a faptelor analizate, subliniind consecvența mai multor agenții în privința acelorași știri. Cu toate acestea, pentru Uscinski (2015), consecvența agențiilor de verificare a faptelor reprezintă pur și simplu o reflecție a prejudecăților politice comune ale verificatorilor care împărtășesc politici partinitoare și accesul la informație. Uscinski (2015) susține că această consecvență a agențiilor de verificare a faptelor indică doar abordarea colectivă a acestora, și nu validarea empirică a consecvenței inter-agenții. Uscinski (2015) pune apoi sub semnul întrebării însuși rolul persoanelor care verifică faptele, dat fiind că astfel de actori nu sunt calificați sau sunt „naivi din punct de vedere epistemologic“ dacă cred că nu e nevoie de calificări și că adevărul este accesibil cu ușurință (p. 247).

Wu, Agarwal, Li, Yang și You (2017) încearcă să depășească unele dintre neajunsurile metodologice ale verificării faptelor prin introducerea metodelor computaționale. Autorii propun metoda „perturbării interogărilor“ ca modalitate de a evita problemele ridicate de selecția subiectivă a faptelor verificate — problemă subliniată de Uscinski (2015) ca neajuns major al verificării faptelor, dat fiind că se bazează pe criterii de selecție nejustificate. Prin perturbarea interogărilor se încearcă extinderea parametrilor faptelor pentru a vedea dacă presupusul adevăr se menține la diferite niveluri de măsurători. Exemplul folosit de Wu et al. (2017) este afirmația fostului primar al orașului New York, Rudy Giuliani, potrivit căreia ratele de adopție au crescut cu 65–70% în perioada în care el a fost la putere. Asta este adevărat conform unei anumite măsurători (1990–1995, 1996–2001 ca serii de date grupate), dar dacă se ajustează la perioada concretă în care Giuliani a fost primar (1994–2001), de fapt a avut loc o scădere cu 1% a ratei adopțiilor. Acest exemplu indică modurile în care datele pot fi manipulate pentru a prezenta adevăruri corecte statistic, dar lipsite de acuratețe contextual.

Abordări computaționale ale studierii informațiilor false

Această secțiune rezumă studiile recente care folosesc abordări computaționale fie pentru a identifica informațiile false, fie pentru a studia răspândirea și efectul acestora.

Majoritatea abordărilor computaționale de detectare a informațiilor false, cum ar fi fake news în social media, folosesc clasificări specifice învățării automate (machine learning) pentru a detecta dacă un articol de știri este adevărat sau nu. Primele abordări automate pentru detectarea informațiilor false pe internet au avut loc în contextul identificării e-mailurilor problematice și a conținutului text problematic de pe website-uri (e.g. spam și farse). Aceste abordări implicau în general aplicarea învățării automate supravegheate asupra textului. Pentru detectarea e-mailurilor de tip spam, de exemplu, asta înseamnă construirea inițială a unui set de date alcătuit din e-mailuri care au fost codate manual ca fiind spam și non-spam, plus un algoritm de clasificare (e.g. regresie logistică, rețea neurală), care este folosit pentru a prezice probabilitatea ca un e-mail să fie problematic pe baza aspectelor evaluate, de exemplu cuvinte-cheie sau tipare ale structurii frazelor.

Castillo, Mendoza și Poblete (2011) au investigat utilizarea metodelor automate de evaluare a credibilității („oferirea de temeuri rezonabile pentru a fi crezute“, p. 675) ale tweet-urilor legate de știri. Mai întâi, o serie de tweet-uri referitoare la știri au fost identificate de un grup de persoane, apoi alt grup de oameni a etichetat aceste tweet-uri ca adevărate sau false, pe baza propriei opinii. Autorii au extras o serie de aspecte ale tweet-urilor legate de abordarea anterioară centrată pe contextul social: aspecte ale postării (e.g. lungimea mesajului, utilizarea de semne de punctuație ca semne de exclamare, impresie pozitivă/negativă exprimată de mesaj, faptul dacă era o postare nouă sau un retweet, dacă conținea sau nu un hashtag); aspecte ale utilizatorului (vechimea contului, numărul de abonați și abonări, număr de tweet-uri scrise); și aspecte de propagare (număr de redistribuiri). S-a

descoperit că informația credibilă are o probabilitate mai mare de a fi răspândită de utilizatori cu conturi mai noi, care sunt mai activi (sub aspectul redistribuirilor), care au mulți follower-i și îi urmăresc, la rândul lor, pe mulți alții. Impresia pozitivă, precum și prezența de semne de întrebare și emoticoane zâmbitoare în tweet-urile care răspândeau știrea au fost asociate cu știri mai puțin credibile, în timp ce prezența de URL-uri în tweet a fost asociată cu o credibilitate mai mare. În ce privește aspectele legate de propagare, este mai probabil ca tweet-urile care au multe redistribuiri să fie considerate credibile.

Tacchini, Ballarin, Della Vedova, Moret și Alfaro (2017) au dezvoltat o abordare de clasificare a postărilor de pe paginile publice de Facebook ca înfățișând teorii valide din punct de vedere științific sau teorii ale conspirației, pe baza „like“-urilor pe care utilizatorii de Facebook le-au dat diferitelor postări.⁷ Abordarea lor se bazează pe presupunerea corelării comportamentului utilizatorilor de Facebook; utilizatorii au tendința de a se grupa prin aprecierea de postări similare, în funcție de preferințele lor de consum al informațiilor referitoare la fapte științifice sau la teorii ale conspirației. Cu toate acestea, după cum au observat autorii, această abordare necesită input din conținutul postării, dat fiind că nu se cunoaște a priori dacă un grup de utilizatori Facebook apreciază aceeași postare pentru că prețuiesc știința sau conspirația.

Vosoughi, Mohsenvand și Roy (2017) au investigat abordările bazate pe învățare supravegheată asupra verificării automate a zvonurilor pe Twitter. Seria de date pe care au utilizat-o a fost constituită din 209 de zvonuri selectate și codate manual referitoare la evenimente reale și autorii s-au concentrat asupra aspectelor legate de stilul lingvistic al tweet-urilor, de utilizatorii care răspândeau zvonurile și de dinamica propagării. Au intenționat să construiască un sistem care putea fi folosit pentru identificarea zvonurilor în timp real și acest sistem a prezis corect veridicitatea a 75% dintre zvonuri mai repede decât sursele publice de încredere (e.g. jurnaliști, oficiali ai forțelor de ordine). Pentru a măsura propagarea, au dezvoltat o abordare centrată pe reconstruirea rețelelor de transmitere în cascadă a tweet-urilor (Twitter API nu oferă acces la cascadele de retweet-uri, ci conectează toate retweet-urile la tweet-ul inițial, indiferent de parcursul concret al mesajelor) folosind orele la care tweet-urile au fost redistribuite și urmărind

relația dintre oamenii care au redistribuit mesajele.⁸ Autorii au descoperit că dinamica temporală ale aspectelor aveau o putere de predicție semnificativă și, dat fiind că aceste dinamici în general nu variază în funcție de mărimea zvonului (e.g. volumul total de atenție pe care a atras-o), acest lucru a permis abordării lor să facă generalizări la evenimente și zvonuri de mărimi diferite.

O descoperire esențială a fost faptul că aspectele de propagare au fost responsabile în cea mai mare parte în ce privește predicția veridicității zvonului și autorii au descoperit că zvonurile care s-au dovedit adevărate aveau tendința de a fi propagate de utilizatori foarte influenți (în termeni de număr de follower-i), care le preluau de la utilizatori cu o influență redusă. O explicație a acestei descoperiri a fost că o persoană cu mare influență nu ar risca să redistribuie un mesaj de la o persoană mai puțin cunoscută decât dacă ar crede că zvonul chiar este adevărat. O altă descoperire esențială a fost că tweet-urile care conțineau zvonuri ulterior descoperite a fi false au avut tendința de a prezenta o distribuție bimodală sub aspectul sofisticării limbajului (în comparație cu alte tweet-uri din baza de date) și cercetătorii au explicat asta prin intenția celor care răspândeau zvonurile; cei care răspândesc zvonuri cu rea-intenție tind să folosească un limbaj sofisticat pentru a acorda zvonului legitimitate și credibilitate, în timp ce persoanele care răspândesc zvonuri fără rea-intenție tind să nu fie atente la limbaj și se exprimă fără sofisticare lingvistică. În sfârșit, autorii au descoperit că zvonurile false tind să fie răspândite de utilizatori influenți și controversați (pentru a stabili cât de controversat este un utilizator se calculează inițial 1 000 de răspunsuri la postările acestuia și apoi se acordă un scor în funcție de proporția dintre răspunsurile pozitive și cele negative, fiind considerați mai controversați utilizatorii cu un raport relativ egal de reacții pozitive și negative), în timp ce persoanele care răspândesc zvonuri ulterior descoperite ca fiind adevărate tind să fie mai influente și mai puțin controversate.

Vosoughi, Roy și Aral (2018) au studiat comportamentul de propagare al știrilor false pe Twitter (după cum am observat anterior, ei au evitat folosirea termenului „fake news”) utilizând abordarea de reconstruire a ramificațiilor cascadelor de retweet-uri (retweet cascade tree) din cercetarea realizată de Vosoughi et al. (2017). Au descoperit că știrile false se răspândesc mai în profunzime (evaluată ca dimensiune a celei mai lungi ramificații a unei cascade de retweet-uri), ajung mai departe (numărul de retweet-uri unice ale

unei știri, reprezentativ pentru numărul de oameni care au consumat știrea) și au o extindere mai mare (dimensiune maximă a unei sub-ramificații). Autorii au descoperit și că știrile false se difuzează mai repede în comparație cu cele adevărate, fiind distribuite către 1500 de oameni de șase ori mai rapid. Aceste tipare au fost pronunțate în special în privința știrilor false din politică, în comparație cu știri referitoare, de exemplu, la terorism, la dezastre naturale sau la alte subiecte.

Autorii au descoperit că aspectele rețelei și utilizatorilor nu explică diferența semnificativă dintre profilurile distribuitorilor de știri adevărate și false. Persoanele care distribuie fake news sunt mai puțin conectate (au mai puțini follower-i și sunt follower-i pentru mai puțini oameni), sunt mai puțin active pe Twitter și conturile lor sunt mai noi; știrile false au o probabilitate de 70% mai mare de a fi redistribuite, în comparație cu cele adevărate, chiar și după controlul acestor factori. Cercetătorii au analizat așadar efectul (asupra difuzării) determinat de o anumită caracteristică a știrilor: noutatea. Au măsurat noutatea știrilor comparând distribuirea temei (identificată prin utilizarea unui model de alocare latentă Dirichlet) știrilor din cadrul eșantionului lor cu tweet-urile la care utilizatorii au fost expuși în ultimele 60 de zile înainte ca ei să redistribuie știrile din cadrul eșantionului. Au descoperit că știrile false sunt mai noi decât știrile adevărate și au sugerat că diferența dintre tiparele de difuzare ar putea fi determinată de faptul că oamenii preferă să distribuie informații noi (indiferent de veridicitatea lor) pentru că noutatea atrage atenția (ne permite să ne actualizăm modul în care înțelegem lumea) și că ar putea exista beneficii sociale interente răspândirii de informație nouă (persoana „este la curent“).

Shin, Jian, Driscoll și Bar (2018) au studiat dinamica procesului de transmitere a informațiilor false în social media urmărind ciclul de viață al 17 zvonuri politice care au circulat pe Twitter în campania prezidențială din SUA din 2012. În concordanță cu descoperirile realizate de Vosoughi et al. (2018), Shin et al. (2018) au descoperit că zvonurile false aveau tendința de a reapărea pe Twitter (adesea suferind „mutații“, i.e. schimbări textuale), prezentând mai multe vârfuri de atragere a atenției, spre deosebire de zvonurile adevărate, care tindeau să aibă o singură perioadă în care se aflau în centrul atenției. Autorii au propus trei motive potențiale pentru aceste tipare temporale diferite. În primul rând, s-ar putea ca persoanele care răspândesc

zvonuri să aibă impresia că zvonurile false trebuie „ajutate“ să ajungă la o mai mare acceptabilitate și sunt astfel necesare încercări repetate de răspândire. În al doilea rând, dat fiind că zvonurile adevărate tind să provină de la instituții consacrate de media, persoanele care răspândesc zvonul s-ar putea să aibă impresia că și-au „epuizat“ potențialul de distribuție și, astfel, nu mai alocă resurse redistribuirii. În sfârșit, răspândirea de zvonuri false s-ar putea realiza în scopul semnalării identității (identity signaling), și nu al răspândirii, încercările repetate de răspândire a zvonurilor false reflectând faptul că persoanele „sunt în aceeași sferă epistemologică“ (p. 285), ceea ce susține legăturile de grup.

Pe lângă examinarea aspectelor de propagare ca aspecte definitorii pentru identificarea fake news, cercetătorii au analizat răspândirea informațiilor false și prin utilizarea de modele formale ale procesului de contagiune. De exemplu, Tambuscio, Ruffo, Flammini și Menczer (2015) au construit modele stocastice de analizare a răspândirii epidemiologice pentru a modela distribuția unei farse ca și cum ar fi un virus. Acestei abordări îi este inerentă perspectiva asupra dezinformării ca simplu proces de contagiune (care nu necesită decât un contact între persoana „infectată“ și cea „neinfectată“ pentru ca transmiterea să aibă loc), însă Törnberg (2018) se bazează pe conceptul de „contagiune complexă“ (conform căruia e posibil să fie necesare mai multe surse pentru a induce efectul), util pentru descrierea răspândirii unor comportamente cum ar fi mișcările sociale și moda avangardistă (Centola și Macy, 2007). Törnberg (2018) ajunge la concluzia că răspândirea de fake news presupune contagiune complexă, dat fiind că decizia unei persoane de a răspândi sau nu poate implica procese de identitate de grup. Modelele de simulare abordează direct contribuția potențială a camerelor de ecou (care, după cum am observat anterior, pot întări identitatea de grup) la răspândirea fake news.

Unii cercetători s-au concentrat asupra rolului pe care îl au programele automate din social media (bots) în răspândirea informațiilor false. De exemplu, Ferrara (2017) a analizat activitatea acestor programe în timpul unei campanii de dezinformare legate de alegerile prezidențiale din Franța din 2017. Contrar așteptărilor, Vosoughi et al. (2018) au descoperit că și programele, și persoanele prezintă comportamente similare pe Twitter în privința răspândirii de știri adevărate și false. Cu toate acestea, chiar dacă

programele automate nu au un efect direct mai mare asupra răspândirii fake news (în comparație cu oamenii), contribuția lor la răspândirea acestora poate influența indirect mai mulți oameni, care la rândul lor vor răspândi fake news. Prin propagarea de neadevăruri, programele automate determină creșterea probabilității ca oamenii să citească știri false, contribuind la percepția că știrile false sunt crezute de mulți oameni, ceea ce ar susține continuarea circulației știrilor false.

Aplicație: cum influențează veridicitatea știrilor dinamice difuzării?

Această secțiune prezintă o analiză a difuzării pe Twitter (prin retweet-uri) a unui eșantion de știri australiene, obiectivul fiind evaluarea modului în care veridicitatea unei știri influențează dinamicele difuzării acesteia.

Retweet-uri și date de rețea

Baza de date este constituită dintr-o colecție de retweet-uri de știri publicate de trei surse media australiene în trei zile selectate aleatoriu din anul 2017 din Ackland, O'Neil și Park (2019).⁹

Pașii de colectare a informațiilor au fost:

(1) Selectarea aleatorie a unei zile din săptămână din trei luni consecutive din 2017 (ne-am asigurat că zilele alese nu au fost dominate de un anumit

eveniment de știri): 22 mai, 16 iunie și 4 iulie.

(2)Colectarea tuturor știrilor difuzate pe Twitter de publicații în acele zile.

(3)Colectarea tuturor retweet-urilor știrilor din următoarele șapte zile.

(4)Colectarea informațiilor referitoare la legăturile din rețea dintre persoanele care au redistribuit știri politice (pe care le-am numit „redistribuitori politici“).

(5)Colectarea informațiilor referitoare la legăturile din rețea dintre redistribuitorii politici și politicieni australieni.

Componenta masivă a rețelei de follower-i pe Twitter (redistribuitori politici și politicieni australieni) este prezentată în Figura 3.1; aceasta este rețeaua în care a fost măsurată difuzarea știrilor (prin retweet-uri). Numărul de noduri este de 495 (12 467) și mărimea nodurilor este proporțională cu gradele interioare. Nodurile din rețea au fost poziționate folosindu-se un algoritm bazat pe linii de forță, în care nodurile conectate apar mai apropiate (ca și cum nodurile ar fi arcuri) și nodurile neconectate sunt mai îndepărtate (ca și cum nodurile sunt încărcate electrostatic și se resping reciproc). Algoritmii bazați pe linii de forță pot dezvălui grupări în rețea, iar aici există dovezi puternice de grupare pe baza ideologiei politice.

Figura 3.1. Rețeaua de follower-i pe Twitter: persoanele care au redistribuit știri politice și politicieni australieni



Notă: Gri-închis — Partidul Laburist (stânga politică), negru — Coaliția Național-Liberală (dreapta politică), gri — alte partide, alb — redistribuitori pe Twitter. Pentru o mai mare claritate a rețelei, nodurile nu sunt afișate.

Selectarea știrilor

Am planificat inițial să selectăm știri care conțineau cel puțin un fapt sau o statistică susceptibilă a fi verificată. Însă un test pilot asupra abordării fact-checking a indicat că erau necesare mai multe criterii de selecție, dat fiind că știrile tradiționale se bazează în mare măsură pe citate și fapte observate, posibilitatea de includere de informații false fiind așadar redusă.

Imparțialitatea în conținutul știrilor este un produs al modificărilor normative care au apărut în secolul XX după cel de-al Doilea Război Mondial. Ca reacție la propaganda din timpul războiului, trusturile de presă și guvernul au încercat să reajusteze imparțialitatea prezentării de știri. Comisia pentru Libertatea Presei (1947) a dezvoltat o normă de știri analitice în care editorialele constituiau rubrici separate în ziare (Metzgar și Hornaday, 2013). Am decis așadar ca verificarea faptelor să se concentreze asupra editorialelor, dat fiind că acestea sunt rubricile în care individualismul și părtinirea sunt de așteptat. În sfârșit, un alt factor a influențat selectarea știrilor ce urmau a fi verificate: dat fiind că intenția noastră era să analizăm difuzarea (prin retweet-uri) știrilor, am ales articole cu un număr mare de redistribuiri pe Twitter.

Au fost analizate în total 14 articole, cinci din Sydney Morning Herald, cinci din Huffington Post și patru din The Australian (Tabelul 3.1). Ne-am concentrat în special asupra editorialelor, deși am inclus și analiza de știri tradiționale pentru a extinde eșantionul, precum și în scopul comparației: opt articole au fost considerate „de opinie” și șase articole au fost clasificate ca „știri tradiționale” (articolele din Huffington Post au fost considerate hibride).

Tabelul 3.1. Știrile din eșantion, scorurile de adevăr ale acestora și nivelurile de difuzare

<i>Numărul știri</i>	<i>Articol</i>	<i>Companie</i>	<i>Data</i>	<i>Format</i>	<i>Scor de adevăr</i>	<i>Răspândire</i>	<i>Viteză</i>	<i>Amploare</i>
1	ALP Gender Equality	Aus	16.6.2017	Articol de opinie	4	1	n.a.	n.a.
2	Turnbull Midwinter Ball	Aus	16.6.2017	Articol de opinie	4	10	2,5	2,67
3	Too Late for Turnbull	Aus	22.5.2017	Articol de opinie	5	13	3,0	2,52
4	Mediscare	Aus	16.6.2017	Articol de opinie	4	1	n.a.	n.a.
5	Scott Morrison Economy	SMH	22.5.2017	Articol de opinie	5	1	n.a.	n.a.
6	Off the Record	SMH	16.6.2017	Articol de opinie	4	1	n.a	n.a.
7	Citizenship Test	SMH	16.6.2017	Articol de opinie	2	8	64,3	2,26
8	Tony Abbott Manifesto	SMH	4.7.2017	Articol de opinie	5	4	3,8	1,57
9	Manus Island	HP	22.5.2017	Articol de opinie/ știre tradițională	4	28	6,3	2,79
10	James Ashby Denial	HP	22.5.2017	Articol de opinie/ știre tradițională	4	12	1,8	2,63
11	Illegal Medicare Machines	HP	4.7.2017	Articol de opinie/ știre tradițională	4	7	1,8	2,45
12	Fake Refugees	HP	22.5.2017	Articol de opinie/ știre tradițională	5	14	4,0	2,00
13	Adani Mine Investment	SMH	22.5.2017	Știre tradițională	5	47	3,5	2,90
14	Manus Settlement	HP	16.6.2017	Știre tradițională	4	4	5,0	1,59

Verificarea știrilor

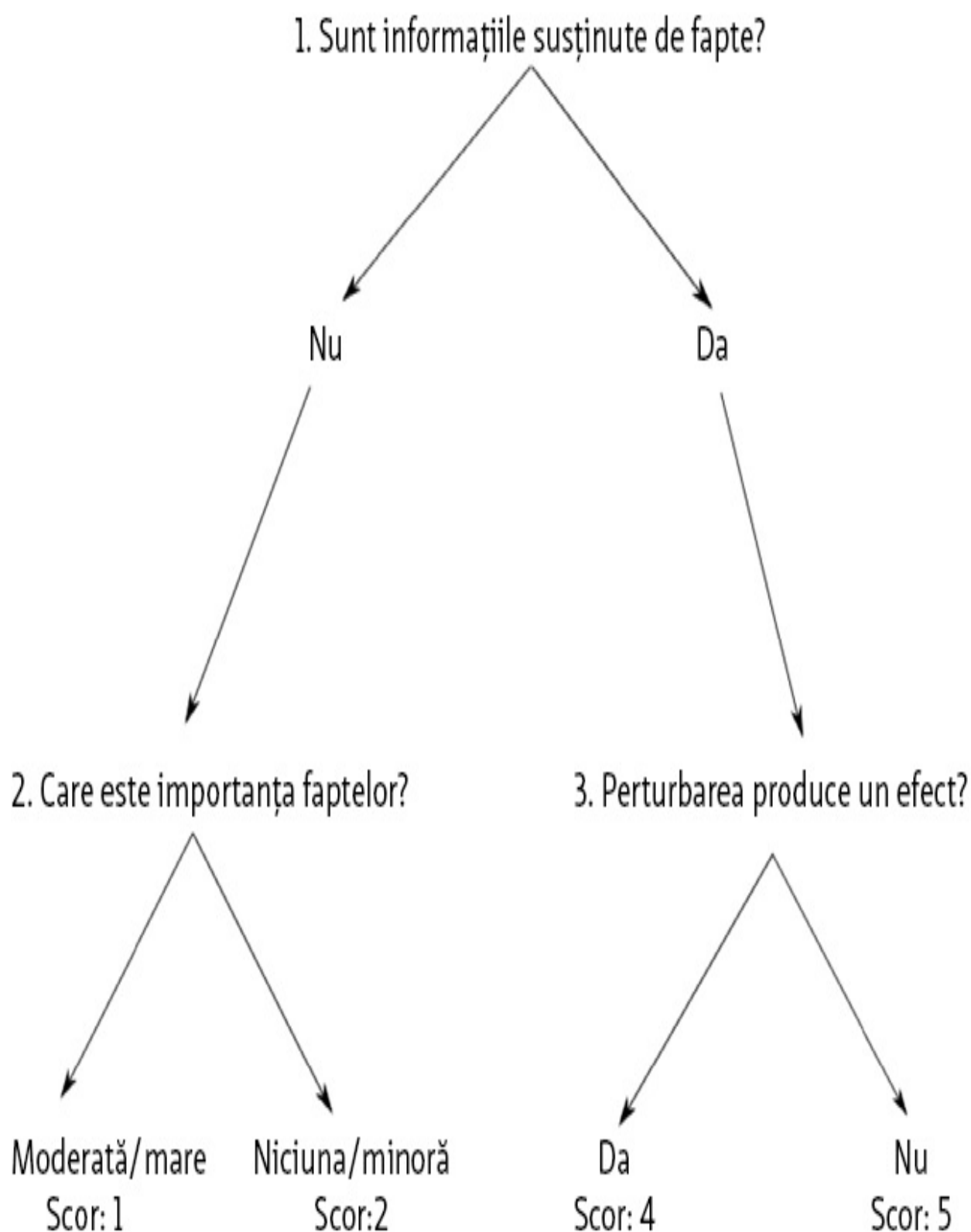
Am dezvoltat și am implementat o metodologie originală de verificare a faptelor prin care am încercat să codăm știrile în funcție de aderența la adevăr. Principalul considerent al metodologiei a fost modul de încadrare a unui model interpretativ de analiză într-un cadru de referință în mare parte cantitativ. Nu este important dacă verificarea faptelor a fost cu adevărat obiectivă sau dacă cel puțin s-a dovedit consecventă. Dar, după cum au arătat Uscinski și Butler (2013), nu se întâmplă așa. Pentru a evita, sau măcar pentru a atenua o parte dintre problemele inerente abordărilor din prezent de verificare a faptelor, am folosit următorul model de lucru (Figura 3.2) pentru a atribui fiecărui articol un scor pe o scală Likert de la 1 (foarte neadevărat) la 5 (foarte adevărat):

(1)Există statistici sau fapte nesusținute de dovezi sau informații din surse de încredere?¹⁰

(2)Care este semnificația statisticilor sau faptelor care nu au fost verificate? Au un efect minor/moderat/mare asupra validității articolului și/sau asupra discursului general al articolului?

(3)În cazul în care au fost verificate fapte sau statistici, perturbări rezonabile ale parametrilor acestora modifică sau compromit intenția sau cadrarea subiacentă a știrii într-un mod consecvent și logic?

Figura 3.2. Procesul de verificare a faptelor



Această metodă include unele elemente fundamentale ale verificării faptelor, cum ar fi verificabilitatea și aderența la adevăr (pasul 1). Nu se poate presupune că verificarea faptelor ar fi o evaluare obiectivă a adevărului, o astfel de perspectivă nefăcând altceva decât să-i diminueze legitimitatea și valoarea. Prima critică pe care Uscinski și Butler (2013) o aduc verificării faptelor este că tinde să promoveze o voce a autorității și certitudinii. Perspectiva persoanei care verifică faptele nu este subliniată îndeajuns, ceea ce sugerează implicit un caracter empiric al verificării faptelor. Deși aceste aspecte sunt considerate uneori ca fiind puncte forte ale verificării faptelor, în contextul acestei cercetări ar fi dus la o reducere a aplicabilității acesteia. Metoda expusă mai sus subliniază limitările sale de interpretare și evită implicarea de adevăruri empirice.

Pașii 2 și 3 încearcă să încorporeze o versiune simplificată a perturbării interogării propusă de Wu et al. (2017). Pasul 2 evaluează efectul afirmațiilor neadevărate prin clasificarea faptelor în funcție de relevanța lor contextuală: dacă argumentul se bazează masiv pe o afirmație neadevărată, scorul este redus. Pasul 3 perturbă parametrii unui fapt pentru a verifica dacă adevărul este rezultatul unei manipulări statistice sau al selectării preferențiale a datelor. Acest pas a fost realizat prin analiza cercetătorilor, și nu prin metode computaționale, astfel încât să se evite comparația directă cu metoda perturbării interogării propusă de Wu et al. (2017), însă este realizat în aceeași idee.¹¹

Modelul de lucru poate fi demonstrat pe scurt utilizând articolul „ALP Gender Equality“ publicat în The Australian în data de 16 iunie 2017. Afirmația „Bill Shorten susține egalitatea de gen între laburiști,[...] dar ideea nu răzbate la nivelul național al partidului“ trebuie să treacă de pasul 1 pentru a fi considerată mai mult decât o critică nejustificată. Articolul menționează apoi că există „dovezi de acțiune în acest sens obținute de The Weekend Australian“: întrucât această afirmație a fost verificată, articolul a trecut de pasul 1 în modelul de lucru. În continuarea articolului se susține că: „Majoritatea rolurilor importante din departamentul condus de Shorten sunt ocupate de bărbați“. Abia la pasul 3 încep faptele din articol să fie mai puțin

convingătoare. Faptele precedente au fost folosite într-un mod care poate fi afectat de perturbări întemeiate. Când comparăm Partidul Laburist cu adversarul său, Coaliția Liberală/Națională, dezechilibrele dintre genuri sunt relativ minore. Articolul își bazează argumentarea pe critica normativă a rolurilor în funcție de gen în cadrul unui partid cu o reprezentare mai bună decât a concurenței. Din acest motiv, articolul nu îndeplinește criteriul pasului 3 și primește un scor total de 4.

Tabelul 3.1 prezintă scorurile de „adevăr” pentru cele 14 articole (verificarea completă a faptelor poate fi oferită de către autori la solicitări directe). Există o aderență relativ mare la adevăr, scorul mediu fiind de 4.2. După cum am anticipat, știrile tradiționale au o aderență mai mare la adevăr în comparație cu editorialele; acestea din urmă sunt adesea scrise de politicieni, care monitorizează informațiile neadevărate în mai mică măsură decât media, reputația instituțiilor de presă depinzând de capacitatea lor de a oferi adevărul. Scorul mediu a fost relativ același în cazul tuturor platformelor de știri: The Australian a avut un scor mediu de 4.3, în timp ce The Sydney Morning Herald și Huffington Post au avut același scor de 4.2. Dată fiind natura interpretativă a analizei, o privire mai atentă asupra excepțiilor poate oferi o imagine mai clară a unei știri care conține informații false într-o măsură semnificativă. Articolul „Testul cetățeniei” a primit cel mai mic scor, 2 din 5, din cauza manipulării clare a faptelor și a contextului, precum și a retoricii senzaționaliste. Pe baza acestui eșantion de articole de știri am ajuns la concluzia că, deși există exemple de informații false răspândite de instituțiile media, acestea nu sunt frecvente, cel puțin nu în instituțiile avute în vedere.

Difuzarea știrilor pe Twitter

După cum am observat anterior, Vosoughi et al. (2018) au reconstruit ramificațiile cascadelor de retweet-uri și apoi au măsurat difuzarea știrilor adevărate și a celor false în aceste ramificații. Noi am utilizat aici o altă

abordare și examinăm difuzarea (prin retweet-uri) știrilor în rețelele de noduri prezentate în Figura 3.1. Nu cunoaștem calea specifică de difuzare a unei știri specifice (fiindcă nu am reconstruit cascada de ramificații ale retweet-urilor), dar știm exact cine anume a redifuzat știrea și când a făcut-o. Figura 3.3 prezintă rețeaua de follower-i, nodurile negre indicând utilizatorii Twitter care au redifuzat știrea „Investiția în mina Adani” (această știre a avut cel mai mare număr de retweet-uri — 47). Se pare că această știre a fost redifuzată în principal de utilizatori Twitter cu înclinații spre stânga politică (urmăresc în general politicieni laburiști), deși există un număr mic de retweet-uri ale știrii de către persoane apropiate de politicieni din Coaliția Național-Liberală (de dreapta).

Figura 3.3. Rețeaua de follower-i pe Twitter — nodurile negre indică utilizatorii care au redistribuit știrea „Investiția în mina Adani”



Am evaluat modul în care adevărul știrii influențează difuzarea sa sub următoarele aspecte:

- (1)Extindere: număr total de retweet-uri ale știrii
- (2)Viteză: număr de ore necesare pentru a se ajunge la 50% din totalul de retweet-uri
- (3)Amplare: distanță geodezică medie sau media celei mai scurte căi calculate pentru toate perechile de persoane care au redifuzat știrea

O provocare a analizei noastre a constatat în faptul că nu avem multe date: nu am codat decât 14 știri, dat fiind consumul mare de timp impus de verificarea faptelor, dar și pentru că am selectat doar știri din trei zile, astfel că am avut un număr limitat de știri care să îndeplinească criteriile noastre de selecție expuse anterior. Mai mult, nu a existat o mare variație a scorului de adevăr al știrilor: toate în afară de una au avut un scor de 4 sau 5.

Cei trei indicatori ai difuzării (extindere, viteză și amplare) sunt prezentați în Tabelul 3.1 și este evident că, în ciuda intențiilor noastre de a selecta știri cu un număr semnificativ de retweet-uri, patru dintre acestea au fost redifuzate doar o dată de utilizatorii Twitter din eșantionul nostru. În realitate, aceste știri au avut un număr mai mare de retweet-uri, dar am inclus în eșantion doar utilizatorii de Twitter care (1) au redifuzat cel puțin una dintre știrile politice și (2) urmăresc sau sunt urmăriți de cel puțin un alt redistribuitor politic sau de un politician australian. Aceste patru știri au fost excluse din analiza următoare.

Pe baza cercetărilor expuse anterior (în special Vosoughi et al., 2018), așteptările noastre erau ca știrile false să aibă o difuzare mai mare (un număr mai mare de retweet-uri), să fie difuzate mai repede și să aibă o amplare mai mare. Datele expuse mai sus arată că nu putem oferi o analiză riguroasă a

tiparelor de difuzare ale știrilor adevărate în comparație cu cele false; intenția noastră este de a ilustra această abordare, și nu de a oferi dovezi concludente. Singura știre care a primit un scor de 2 (neadevărată) — „Test de cetățenie“ — a avut o extindere de opt retweet-uri (mai mică decât media tuturor știrilor, de 15 retweet-uri), a durat 64 de ore pentru a ajunge la 50% din totalul numărului de retweet-uri (mult mai mult decât media de zece ore) și până la finalul acestei perioade, lungimea medie a căii celei mai scurte (în rețeaua de follower-i) dintre oamenii care au redifuzat știrea a fost de 2,26 (mai mică decât amplitudinea medie a totalului de știri, care a fost 2,34). Așadar, conform celor trei măsurători, tiparul de difuzare al știrii neadevărate din eșantion, în comparație cu media, a fost contrar așteptărilor noastre. Cu toate acestea, după cum am observat deja, limitările impuse de date ne permit prezentarea aplicației noastre ca o contribuție metodologică, și nu ca test definitiv al comportamentului de difuzare a știrilor adevărate și false.

Concluzie

În acest capitol am trecut în revistă cercetările de ultimă oră din domeniul informațiilor false și al social media, concentrându-ne asupra abordărilor al căror scop este identificarea informațiilor false (și a celei mai comune forme pe care o îmbracă acestea în zilele noastre, fake news), precum și cercetările menite a ne aprofunda cunoașterea modului în care informațiile false se răspândesc și a consecințelor acestora. Deși ne-am concentrat asupra abordărilor computaționale care se folosesc de datele din social media, am expus și ultimele cercetări asupra metodei de verificare manuală a știrilor.

Am oferit apoi o aplicație empirică al cărei scop a fost caracterizarea tiparelor de difuzare a știrilor adevărate și false din politica australiană. Aplicația a demonstrat o nouă abordare de verificare a faptelor din editoriale și am încercat, de asemenea, să demonstrăm că e posibilă studierea difuzării știrilor chiar și în absența cascadelor de ramificații ale retweet-urilor (care nu sunt oferite de Twitter și a căror reconstruire reprezintă o provocare). Cu toate

astea, aplicația noastră empirică a întâmpinat o provocare în faptul că am avut puține date (știri codate statistic) și că nu a existat o variație semnificativă a scorurilor de adevăr. Oferim aplicația noastră empirică drept demonstrație a unei noi abordări a studierii fake news utilizând datele Twitter.

Note

1.Unii autori fac diferența între misinformation și disinformation, prima implicând un actor care răspândește informații inexacte pe care le consideră adevărate, iar a doua implicând o încercare intenționată de a induce în eroare. În practică, disinformation este tratată în general ca subcategorie a misinformation, dat fiind că e mai dificilă evaluarea motivelor pentru care actorii transmit informații inexacte.

2.Observând politizarea din prezent a sintagmei „fake news“, Vosoughi et al. (2018) preferă termenul „false news“ („știri false“).

3.Bakshy, Messing și Adamic (2015) compară efectele relative ale camerelor de ecou și ale bulelor filtrante asupra expunerii unei persoane la diferite conținuturi politice pe Facebook.

4.Din câte știm, nu s-a testat dacă susținătorii lui Trump aveau o tendință mai mare de a consuma fake news sau dacă pur și simplu erau expuși la un volum mai mare de fake news consecvente propriilor atitudini.

5.O serie temporală X este considerată a constitui „cauzalitate Granger“ pentru o serie Y dacă valorile lag ale lui X sunt semnificative statistic într-o regresie care prezice valorile viitoare ale lui Y .

6.Informațiile următoarele sunt preluate din Shu et al. (2017).

7.Vosoughi et al. (2018) susțin că cercetările răspândirii studiilor științifice și ale teoriilor conspirației nu permit înțelegerea factorilor care influențează

răspândirea știrilor adevărate și a celor false, dat fiind că o teorie a conspirației este prin definiție neverificabilă.

8. Autorii observă că și Castillo et al. (2011), și Kwon, Cha, Jung, Chen și Wang (2013) examinează utilizarea aspectelor propagandistice pentru predicția veridicității tweet-urilor, dar acești autori nu identifică ramificațiile cascadei de tweet-uri, astfel că nu au putut să folosească aspecte propagandistice la fel de sofisticate precum cele folosite de Vosoughi et al. (2017).

9. Cele trei surse media australiene au fost The Australian, Sydney Morning Herald și Huffington Post. Au fost alese în scopul de a acoperi un spectru larg de afilieri politice, The Australian situându-se la dreapta politică, Sydney Morning Herald, la stânga și Huffington Post, potențial în orice parte.

10. Sursele de încredere sunt dovezile care provin fie dintr-o sursă primară, fie dintr-o sursă secundară a cărei integritate trebuie stabilită în funcție de probe istorice ale veridicității. Am evaluat acest aspect pe baza statutului sursei, determinat prin indicatori externi, ca veridicitatea dovedită în trecut, responsabilitatea și corectări retrospective. Verificarea faptelor prin această metodologie trebuie să evite pe cât posibil orice afiliere și distorsiune. Acestea sunt imposibil de eradicat complet, așa că verificarea faptelor trebuie considerată o evaluare subiectivă a ceea ce e adevărat sau fals.

11. Codarea a fost realizată de unul dintre coautori, dar într-un studiu la scală mai mare ar fi preferabil să existe mai multe persoane care se ocupă de verificarea faptelor și apoi să se calculeze fidelitatea interevaluatori.

Referințe

Ackland, R., O'Neil, M. și Park, S. (2019). Engagement with news on Twitter: Insights from Australia and Korea. *Asian Journal of Communication*, 29 (3), 235–251.

Allcott, H. și Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31 (2), 211–235.

Amazeen, M.A. (2015). Revisiting the epistemology of fact-checking. *Critical Review*, 27 (1), 1–22.

Bakshy, E., Messing, S. și Adamic, L.A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348 (6239), 1130–1132.

Castillo, C., Mendoza, M., și Poblete, B. (2011). Information credibility on twitter. *Proceedings of the 20th International Conference on World Wide Web (WWW)*, Hyderabad, India.

Centola, D. și Macy, M.W. (2007). Complex contagion and the weakness of long ties. *American Journal of Sociology*, 113 (3), 702–734.

Ecker, U.K., Lewandowsky, S., Fenton, O., și Martin, K. (2014). Do people keep believing because they want to? Preexisting attitudes and the continued influence of misinformation. *Memory & Cognition*, 42 (2), 292–304.

Ferrara, E. (2017). Disinformation and social bot operations in the run up to the 2017 French presidential election. *First Monday*, 22 (8).

Guess, A., Nyhan, B., și Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign. Accesat la www.ask-force.org/web/Fundamentalists/Guess-Selective-Exposure-to-Misinformation-Evidence-Presidential-Campaign-2018.pdf

Hasher, L., Goldstein, D. și Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning & Verbal Behavior*, 16, 107–112.

Katz, E. și Lazarsfeld, P.F. (1955). *Personal influence: The part played by people in the flow of mass communications*. Glencoe, IL: Free Press.

Kwon, S., Cha, M., Jung, K., Chen, W. și Wang, Y. (2013). Prominent features of rumor propagation in online social media. *Proceedings of the 13th*

International Conference on Data Mining (ICDM), IEEE, 1103–1108.

Leetaru, K., și Schrod, P.A. (2013). GDELT: Global data on events, location, and tone, 1979–2012. International Studies Association Meetings, San Francisco.

Lewandowsky, S., Ecker, U.K., Seifert, C.M., Schwarz, N., și Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13 (3), 106–131.

Metzgar, E.T. și Hornaday, B.W. (2013). Leaving it there? The Hutchins commission and modern American journalism. *Journal of Mass Media Ethics*, 28 (4), 255–270.

Nyhan, B. și Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32 (2), 303–330.

Shin, J., Jian, L., Driscoll, K. și Bar, F. (2018). The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior*, 83, 278–287.

Shu, K., Sliva, A., Wang, S., Tang, J. și Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19 (1), 22–36.

Tacchini, E., Ballarin, G., Della Vedova, M.L., Moret, S. și Alfaro, L. (2017). Some like it Hoax: Automated fake news detection in social networks. *Proceedings of the Second Workshop on Data Science for Social Good*, Skopje, FYR Macedonia. CEUR Workshop Proceedings, Volume 1960.

Tajfel, H. și Turner, J.C. (2001). An integrative theory of intergroup conflict. În M.A. Hogg și D. Abrams (coord.), *Key readings in social psychology: Intergroup relations: Essential readings* (pp. 94–109). New York, NY, US: Psychology Press.

Tajfel, H. și Turner, J.C. (2004). The social identity theory of intergroup behavior. În J.T. Jost și J. Sidanius (coord.), *Key readings in social*

psychology: Political psychology: Key readings (pp. 276–293). New York, NY, US: Psychology Press.

Tambuscio, M., Ruffo, G., Flammini, A. și Menczer, F. (2015). Fact-checking effect on viral Hoaxes: A model of misinformation spread in social networks, WWW 2015 Companion.

Törnberg, P. (2018). Echo chambers and viral misinformation: Modeling fake news as complex contagion. PLoS One, 13 (9).

Uscinski, J.E. (2015). The epistemology of fact checking (is still naïve): Rejoinder to Amazeen. Critical Review, 27 (2), 243–252.

Uscinski, J.E. și Butler, R.W. (2013). The epistemology of fact checking. Critical Review, 25 (2), 162–180.

Vargo, C.J., Guo, L. și Amazeen, M.A. (2018). The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. New Media & Society, 20 (5), 2028–2049.

Vosoughi, S., Mohsenvand, M.N. și Roy, D. (2017). Rumor Gauge: Predicting the veracity of rumors on Twitter. ACM Transactions on Knowledge Discovery from Data (TKDD), 11 (4), 50.

Vosoughi, S., Roy, D. și Aral, S. (2018). The spread of true and false news online. Science, 359 (6380), 1146–1151.

Wu, Y., Agarwal, P.K., Li, C., Yang, J. și Yu, C. (2017). Computational fact checking through query perturbations'. ACM Transactions on Database Systems, 42 (1), 1–41.

Young, D.G., Jamieson, K.H., Poulsen, S. și Goldring, A. (2018). Fact-checking effectiveness as a function of format and tone: Evaluating FactCheck.org and FlackCheck. org. Journalism & Mass Communication Quarterly, 95 (1), 49–75.

Capitolul 4. Articole retractate — versiunea științifică a fake news

Judit Bar-Ilan și Gali Halevi

Introducere

Progresele științifice se fac treptat. „Am stat pe umerii gigantilor“ este un citat bine-cunoscut din Newton¹ și se referă la procesul științific prin care fiecare articol care expune noi informații îi permite comunității științifice să ajungă la o mai mare cunoaștere pe baza descoperirilor de până atunci. Scopul publicațiilor științifice a fost întotdeauna de a reprezenta o sursă de încredere, validată, de date care pot fi împărtășite și amplificate. Ca parte a acestei responsabilități, publicațiile științifice impun măsuri prin care se asigură că acestea au un conținut într-adevăr original, de încredere, și că rezultatele cercetărilor pot fi reproduse, menținând astfel intactă integritatea științifică. Unul dintre principalele moduri prin care comunitatea publicațiilor științifice asigură veridicitatea științifică este procesul de evaluare inter pares, practică prin care fiecare articol trimis unei reviste este analizat de cel puțin doi alți specialiști din același domeniu, întotdeauna anonimi și care adesea nu știu cine sunt autorii. Aceste măsuri sunt luate tocmai pentru a asigura o evaluare nepărtinitoare și centrată pe articole și pe descoperirile acestora. Principala sarcină a evaluatorilor este să verifice originalitatea, calitatea și integritatea fiecărui articol și să examineze acuratețea datelor, analizei, rezultatelor, concluziilor ș.a.m.d. Deși evaluatorii pot recomanda modificări majore sau minore, au și puterea de a respinge articole pe care le găsesc ca nefiind de încredere din diferite motive (Biagioli, Kenney, Martin

și Walsh, 2018; Bozzo, Bali, Evaniew și Ghert, 2017; Nicholas et al., 2015).

Cu toate acestea, în ciuda acestui proces riguros, se întâmplă frecvent ca articolele să fie retrase din reviste din diferite motive (Budd, Sievert, Schultz și Scoville, 1999; Cokol, Ozbay și Rodriguez-Esteban, 2008; Fang, Steen și Casadevall, 2012; Steen, 2010, 2011). Uneori motivul retragerii este evident și ușor de corectat (Bar-Ilan și Halevi, 2018; Budd, Sievert și Schultz, 1998; Halevi și Bar-Ilan, 2016; Williams și Wager, 2013). Aceste tipuri de retrageri au loc din cauza erorilor administrative, a greșelilor de citare sau a greșelilor de revizuire care nu au fost identificate suficient de devreme pentru a fi corectate înainte de publicare, articolul fiind astfel retras cu un aviz de retractare. În majoritatea cazurilor, aceste articole sunt corectate și republicate. Însă există motive mult mai grave de retractare, în special cele provocate de falsificarea datelor, de manipularea rezultatelor, de abordări neetice, de plagiat etc. (Almeida, de Albuquerque Rocha, Catelani, Fontes-Pereira și Vasconcelos, 2015; Corbyn, 2012; Fang și Casadevall, 2011; Inoue și Muto, 2016; Noyori și Richmond, 2013).

În domeniile medical și biomedical, retractarea articolelor provocată de probleme etice și de existența datelor, analizelor și descoperirilor eronate sau chiar inventate trebuie examinată foarte atent, dat fiind efectul negativ enorm pe care îl au asupra viitoarelor practici medicale și, mai important, asupra vieților omenești. În acest capitol ne concentrăm asupra articolelor retractate în domeniul medical și biomedical și demonstrăm modul în care au afectat și, în unele cazuri, continuă să afecteze comunitatea medicală și societatea. Pentru a ilustra gravitatea articolelor retractate din domeniul medical și potențialul de a expune pacienții la riscuri, un studiu realizat de G. Steen (2011) a evaluat peste 180 de cercetări primare retrase și 851 de cercetări secundare care descriu rezultatele mai multor studii la care au luat parte în total peste 28 000 de persoane. Cercetările primare au fost considerate cele care au fost inițial retrase, iar cercetările secundare au fost cele care citau articolele retractate, bazându-se pe acestea pentru colectarea de date, metode sau analiză. Examinarea a arătat că 70 de articole care au fost retractate din cauza fraudei au constatat în studii care au tratat mai mulți pacienți decât cele retractate din cauza erorilor. Autorul a descoperit și că aceste articole au fost citate de peste 5 000 de ori și că citările au fost în cea mai mare parte referitoare la cercetare, ceea ce înseamnă că ideile exprimate în articolele

frauduloase au influențat cercetări ulterioare și pacienții care au participat la studii. Așadar, informația incorectă sau frauduloasă raportată în studii și reviste medicale poate afecta pacienții în mod direct prin continuarea citărilor, prin preluarea metodelor și prin aplicarea tratamentelor sau a ideilor din publicația incriminată.

Există mai multe cazuri foarte celebre de astfel de retractări care au ajuns subiect de știri și au provocat neliniștea publicului prin gravitatea lor. Printre acestea a fost articolul din 1998 de dr. Wakefield din Lancet, care sugera că vaccinul combinat pentru pojar, oreion și rubeolă poate provoca autism la copii. Acest articol — retractat în 2010 din cauza metodelor deficitare și a conflictelor financiare ale autorului — rămâne însă până în ziua de azi foarte influent, mulți părinți și multe grupuri anti-vaccin fiind reticente în privința vaccinării copiilor, ceea ce a provocat apariția mai multor cazuri de boală (Facts about the Measles Outbreak, 2015). Alte cazuri celebre sunt reprezentate de dr. Drasee, cercetător în cardiologie de la Facultatea de Medicină Harvard, care a fabricat cea mai mare parte a cercetărilor sale, peste 100 de articole fiind retractate din cauza datelor frauduloase (Broad, 1983) sau dr. Straus, cercetător în oncologie care a fost acuzat că a falsificat rezultatele cercetării prin raportarea de pacienți neînscriși, prin administrarea de doze falsificate de medicamente (Upi, 1982) și prin nerespectarea regulilor de consimțământ.

Pe lângă aceste cazuri bine cunoscute, retractări în domeniile medical și biomedical au loc cu regularitate și, deși în mare nu pun viața în pericol, faptul că sunt citate în continuare în literatura medicală și că sunt distribuite ca fiind cercetări legitime poate reprezenta o amenințare la adresa sănătății chiar și după ani întregi de când au fost retrase.

Sursele și colectarea datelor

Pentru a examina caracteristicile și efectul articolelor medicale și biomedicale

retractate am folosit Retraction Watch Database („Retraction Watch Database“, fără dată), curatoriată de Adam Marcus și Ivan Oransky, proprietarii „Retraction Watch“ („Retraction Watch Database“, fără dată), un blog dedicat înregistrării și semnalizării de articole retractate din diferite discipline științifice. Date fiind cercetările noastre anterioare asupra articolelor retractate, am avut acces la datele curatoriate din baza de date, care include peste 17 000 de articole retractate, notificări de retractare și motive ale retractării. Datele includ și metadate valoroase, inclusiv numele autorului/autorilor, titlul revistei, editura, anul și disciplina. Datorită clasificării în funcție de disciplină din baza de date, am putut să accesăm toate articolele medicale și biomedicale retractate care au fost publicate între 2010 și 2014 și au fost retractate până în 2017. Datele noastre includ 1 294 de cercetări din aceste domenii.

Principalul motiv pentru care am făcut această selecție a fost scopul nostru de a identifica articolele recente retractate din aceste domenii pentru a putea să monitorizăm atenția pe care au primit-o în social media. Utilizarea de social media pentru distribuirea, discutarea sau promovarea cercetării este un fenomen recent. În zilele de început ale social media, utilizarea acestora se desfășura în cea mai mare parte în scopuri personale. Însă, în ultimii zece ani, din ce în ce mai mulți oameni de știință folosesc platforme ca Twitter și Instagram pentru a-și promova cercetarea, pentru a socializa cu alții din același domeniu și pentru a discuta diferite subiecte folosindu-se de aceste instrumente. Unul dintre principalele noastre interese a fost ciclul de viață al articolelor retractate în social media, dat fiind că astfel de discuții nu rămân în sfera științifică — implică nespecialiști care sunt follower-i ai oamenilor de știință sau medicilor și care iau parte activ la distribuirea și discutarea cercetărilor. Așadar, pentru a putea să examinăm indicatorii social media, am selectat articole publicate când social media erau utilizate pe scară largă de către oamenii de știință. Am colectat și numărători ale citărilor, acestea fiind indicatori de impact în cea mai mare parte științific. Toate datele, inclusiv citările, au fost colectate în februarie 2019 pentru a ne asigura că a trecut suficient timp după retractare pentru ca citările să apară în literatura științifică și pentru a ne asigura că pe diferitele platforme apar și alți indicatori, cum ar fi audiența și mențiunile în social media.

Pentru a evalua efectul și în viața științifică, și în cea publică a articolelor

retractate, am folosit mai multe instrumente pentru a număra citările de dinainte și de după retractare, mențiunile în social media și audiența articolelor retractate. Datele referitoare la citări, la mențiunile din social media și la audiență au fost colectate în februarie 2019. Dat fiind că de obicei trece o perioadă de timp între data publicării și data retractării, este importantă examinarea citărilor și atenției primite de articol înainte și după retractare. Deși citările științifice ale articolelor înainte de a fi retractate constituie un fenomen acceptabil, astfel de citări post-retractare sunt îngrijorătoare și ridică întrebarea: de ce continuă oamenii de știință să citeze un articol retractat? În unele cazuri, citarea este negativă, adică autorii citează un articol retractat pentru a da un exemplu de cercetare slabă sau pentru a sublinia rezultatele care le contrazic pe cele din articolul retractat. Însă vedem adesea și citări valide, pozitive, care ingoră complet faptul că articolul a fost retras din literatură (Bar-Ilan și Halevi, 2017). Deși citările sunt considerate a avea un impact științific, audiența, de exemplu, poate fi un indicator al impactului atât științific, cât și social, dat fiind că nu se poate ști cine anume citește aceste articole. În ciuda faptului că publicațiile retrag articole și oferă notificări în acest sens, versiuni anterioare ale articolelor pot fi găsite în baze de date și sunt disponibile gratuit. Aceste versiuni nu includ notificarea de retractare și cititorii s-ar putea să nu cunoască defectele care au provocat retractarea. În plus, am descoperit și că articolele retractate ajung să fie oferite gratuit pe platformele editurilor (Bar-Ilan și Halevi, 2017; Bar-Ilan și Halevi, 2018), ceea ce înseamnă că sunt gratuite pentru persoane care nu trebuie să plătească pentru a le citi. În ciuda faptului că sunt clar marcate ca „Retractate“, gratuitatea acestor articole înseamnă că vor continua să fie citite și distribuite post-retractare, veridicitatea acestui conținut reprezentând o problemă serioasă.

Spre deosebire de citări, distribuirea și mențiunile pe social media sunt reprezentative și pentru impactul pe care cercetarea științifică îl are asupra publicului. Conform unei cercetări realizate de Pew, Facebook și Twitter devin rapid surse semnificative de informații științifice și medicale (Greenwood, Perrin și Duggan, 2016). Așadar, articolele falsificate, neadevărate, devin o amenințare reală la adresa sănătății publice atunci când sunt distribuite online prin rețele de socializare și uneori prin instituții de presă. Principala problemă a distribuirii pe social media este că aceste articole sunt mult mai dificil de identificat și de corectat. Odată ce un articol ajunge în

centrul atenției pozitive sau negative, s-ar putea ca o retractare formală să nu reușească să schimbe percepția oamenilor în același mod în care o face în comunitatea științifică. Deși instituțiile de presă pot aduce corectări și pot anunța că o anumită cercetare nu e „bună” după publicare, se poate ca asta să nu se întâmple suficient de repede sau într-o modalitate care va determina schimbarea percepției publice.

Pentru a evalua ambele fațete — impactul științific și în viața publică al articolelor retractate, am folosit baza de citări Scopus pentru a identifica citările înainte și după retractare. Scopus este o bază de date amplă, care indexează milioane de articole din numeroase discipline științifice. Principalul avantaj al utilizării Scopus pentru identificarea citărilor este că metadatele oferă indicații ale citărilor în funcție de ani, permițându-ne astfel să vedem citările de dinainte și de după retractare. Aceasta este o piesă importantă de puzzle, pentru că ne permite să aflăm de câte ori este citat un articol după ce a fost eliminat din literatură.

Pentru a examina atenția socială și mediatică acordată articolelor retractate, precum și audiența acestora, am folosit platforma PlumX și datele de la Altmetric.com — ambele agregatoare de indicatori altmetrics. Altmetrics este un termen general utilizat pentru a descrie „evaluări alternative” în știință. Evaluările tradiționale sunt reprezentate de citări, iar altmetrics includ alți indicatori, cum ar fi audiența, numărul de descărcări, menționări în social media și altele. Agregatorii identifică menționările articolelor în diferite social media, cum ar fi Twitter și Facebook, bloguri ș.a.m.d. În plus, identifică articolele care au fost menționate în presa de știri și numără și accesările pe Reddit și alte siteuri. Pe lângă aceasta, am evaluat numărul de cititori și pe Mendeley, renumită rețea științifică unde cercetătorii pot salva și distribui articole științifice. Agregatoarele de date identifică menționările prin identificatori, în principal prin link-ul Digital Object Identifier (DOI). PlumX afișează aceste evaluări pentru fiecare articol și prezintă comentarii și articole de știri și de blog. Altmetric oferă link-uri către bloguri și afișează tweet-uri, ceea ce ne-a permis să examinăm tonul și conținutul acestor mențiuni pentru o mică subcategorie de articole retractate, permițându-ne așadar să evaluăm dacă post-retractare impactul acestora a fost persistent negativ sau pozitiv. Audiența Mendeley a fost preluată direct de la Mendeley. Pentru colectarea de date de la Mendeley și Altmetric.com am folosit instrumentul gratuit creat

de Mike Thelwall, The Webometric Analyst (<http://lexiurl.wlv.ac.uk/>).

Câteva aspecte

De unde vin retractările?

Pentru a identifica originea geografică a articolelor retractate ne-am raportat la țara de afiliere a autorilor. Fiecare articol publicat trebuie să conțină numele și țara de afiliere a autorului. Această informație a fost utilizată în analiza noastră pentru a identifica țările autorilor articolelor retractate. Trebuie să menționăm că fiecare articol din baza noastră de date de 1 294 de cercetări are mai mulți autori. Marea majoritate a articolelor științifice sunt publicate în colaborare, echipe de cercetători lucrând împreună și publicând rezultatele împreună. De multe ori aceste colaborări sunt internaționale și autorii sunt din mai multe țări. Am creat o hartă cu toate țările autorilor acestor articole; astfel, un articol din baza de date care are mai mulți autori din țări diferite este trecut în toate aceste țări. După cum se poate vedea pe hartă (Figura 4.1), majoritatea articolelor retractate din baza noastră de date provin din SUA (341), China (309), India (115) și Japonia (68).

Figura 4.1. Originile geografice ale articolelor retractate (prin Tableau)



Sursa hărții: creață de autori

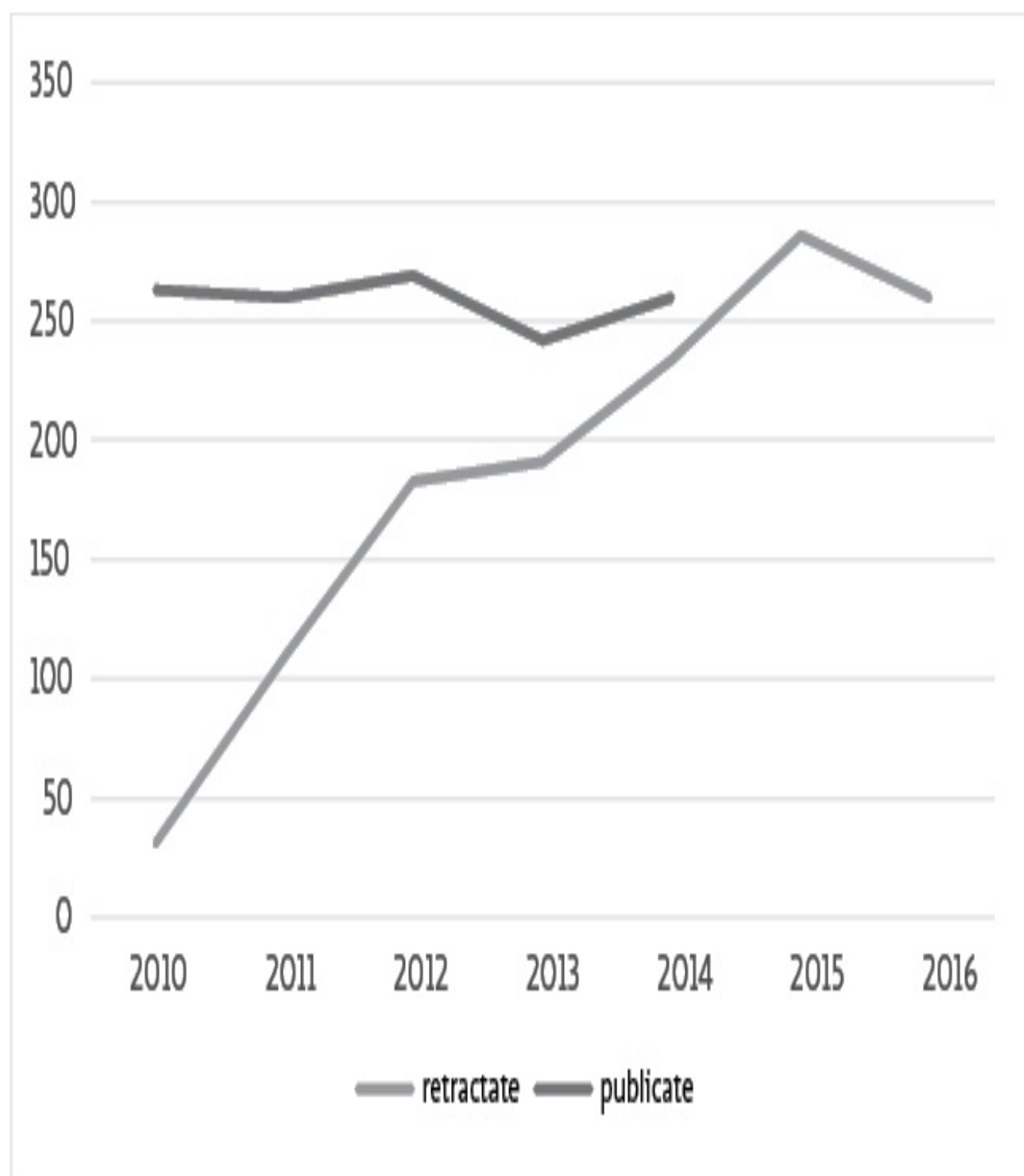
Unul dintre avantajele oferite de Retraction Watch Data Base este că expune și motivele retractărilor și le include în peste 70 de categorii. În majoritatea cazurilor, fiecărui articol îi sunt alocate mai multe categorii. Analiza noastră a arătat că articolele retractate din SUA au fost retrase în cea mai mare parte din cauza copierii/manipulării de imagini, precum și a falsificării/inventării datelor. Articolele chinezești au fost retractate în cea mai mare parte din cauza evaluărilor inter pares false, descoperite printr-o investigare de către revistă sau din cauza plagiatului. Articolele indiene au fost retractate în cea mai mare parte din cauza plagiatului, iar articolele japoneze, din cauza rezultatelor îndoielnice și a inventării/falsificării de date. Un articol interesant, publicat în 2017 (Fanelli, Costas, Fang, Casadevall și Bik, 2017) oferă o perspectivă de ansamblu asupra principalelor motive pentru care oamenii de știință inventează sau falsifică date sau apelează la tertipuri. Primul motiv este presiunea de a publica; oamenii de știință sunt extrem de presați să producă un anumit număr de articole în fiecare an pentru a-și păstra granturile sau pentru a primi o catedră sau o promovare. Așa stau lucrurile în special în SUA și în Marea Britanie, unde numărul de publicații și numărul de citări reprezintă principalii indicatori prin care se evaluează performanța științifică. Lucrurile sunt și mai complicate în țări precum China, unde oamenii de știință sunt răsplătiți cu bani pentru publicarea unui anumit număr de articole. Acești factori, la care se adaugă faptul că există diferențe între țări în privința definirii și sancționării abaterilor, creează medii în care abaterile sunt mai frecvente (Redman și Merz, 2008).

Publicații și retractări în funcție de an

O descoperire interesantă pe care am observat-o în analiza noastră a fost raportul dintre numărul relativ de publicații și numărul retractărilor. Deși, din 2010 încoace, numărul anual al publicațiilor care sunt retrase în anii următori rămâne stabil, aproximativ 260 de publicații pe an (în baza noastră de date),

numărul retractărilor articolelor publicate anterior crește în fiecare an (vezi Figura 4.2).

Figura 4.2. Publicarea anuală de articole retractate și retractări



Sursă: creată de autori

O parte din motivul creșterii numărului anual de retractări poate fi atribuit faptului că, în mediul conectat din ziua de azi, acuratețea datelor și plagiatul pot fi verificate mult mai repede, iar erorile pot fi comunicate mai ușor editorilor, care pot posta notificări de retractare pe site-urile revistelor. Acest proces mai rapid stă la baza reducerii timpului necesar pentru retractarea articolelor (vezi Tabelul 4.1).

Tabelul 4.1. Media numărului de ani necesari pentru retractare în funcție de anul publicării

<i>An</i>	<i>Număr de publicații</i>	<i>Media numărului de ani necesari pentru retractare</i>
2010	263	2,67
2011	260	2,15
2012	269	1,87
2013	242	1,51
2014	260	1,07
Totți anii	1294	1,86

Am evaluat și intervalul de timp dintre publicare și retractare și am calculat media numărului de ani. După cum se vede în Tabelul 4.1, intervalul descrește. În 2010 era necesară în medie o perioadă de 2,7 ani, iar în 2014 ajunsese la 1,1 an. Această descoperire este foarte încurajatoare și arată că cercetările defectuoase sunt scoase din circulație mult mai repede decât în trecut. În domeniile medical și biomedical, unde descoperirile pot influența studii ulterioare imediat ce sunt raportate, este esențială eliminarea cercetărilor frauduloase imediat ce este posibil, astfel încât să se evite prejudiciile asupra pacienților sau publicului general.

Motive ale retractărilor

Am creat un nor de cuvinte-cheie (word cloud) din textul „motive de retractare”, unul dintre câmpurile oferite de Retraction Watch. Într-un nor de cuvinte-cheie, dimensiunea cuvântului sau sintagmei este determinată de numărul de incidențe. Cu cât apare de mai multe ori, cu atât e mai mare. În norul de cuvinte-cheie prezentat în Figura 4.3 este evident că cele mai frecvente motive de retractare a articolelor medicale și biomedicale din baza noastră de date sunt investigările întreprinse de revistă/editură. O investigație nu este în sine un motiv de retractare, dar, uitându-ne la expresiile din norul de cuvinte-cheie, vedem că motivele dominante sunt plagiatul, manipularea imaginilor, datele eronate, abateri ale autorului, falsificarea datelor și false evaluări inter pares. Dominante sunt și reproducerea articolului și a imaginilor, ceea ce indică în general faptul că autorii și-au reciclat articolele, trimițând același material mai multor reviste. Când trimite un articol către o revistă, una dintre întrebările la care un autor trebuie să răspundă și unul dintre lucrurile pe care trebuie să le ateste este originalitatea materialului, ceea ce înseamnă că articolul nu a mai fost trimis în altă parte. Un autor care își reciclează materialele pentru a submina procesul de publicare încalcă deliberat regulile etice ale publicării de articole științifice.

Sursă: Creat de autori

Notă: ORI — United States Office of Research Integrity. (N.r.)

Figura 4.3. Norul de cuvinte-cheie (word cloud) cu motivele retractării

Îndoieli referitoare la originalitate
Îndoieli referitoare la rezultate
Rezultate nereproductibile
Articol plagiat
Text plagiat
Eroare a revistei/editurii
Manipulare a imaginilor
Date eronate
Date îndoielnice
Eufemism pentru duplicare
Abateri ale autorului
Îndoieli referitoare la date
Fals autor
Erori de analiză
Investigații de către ORI
Investigații de revistă/editură
Imagine îndoielnică
Metode eronate
Duplicare a articolului
Duplicare a datelor
Încălcarea reglementărilor de către autor
Duplicarea textului
Obiecții aduse de terți
Absența acordului autorului
Reclamație pentru nerespectarea drepturilor de autor
Falsă evaluare *inter pares*
Falsificarea/inventarea imaginilor
Imagini eronate
Investigații întreprinse de terți
Falsificarea/inventarea datelor
Rezultate și/sau concluzii eronate
Încălări etice de către autor
Notificare — informații limitate sau absente

Un alt aspect interesant al acestui nor de cuvinte-cheie este numărul de retractări din cauza problemelor legate de date. Să observăm mărimea sintagmelor ca „inventarea datelor“, „date eronate“, „îndoieli referitoare la date“, „probleme de date“, „duplicarea datelor“ și altele. Asta înseamnă că există un număr mare de articole care au fost eliminate din literatura științifică din cauza unor probleme grave la nivelul datelor raportate. Deși în alte domenii ale comunicării științifice problemele legate de date sunt grave, în domeniile medical și biomedical aceste probleme sunt mai severe, în principal pentru că cercetarea în aceste discipline are legătură directă cu tratamentele și procedurile medicale și cu medicamentele prescrise pacienților. În plus, datele medicale falsificate pot induce publicul în eroare și le pot provoca oamenilor vătămări reale.

Reviste importante

Tabelul 4.2. Lista revistelor cu cele mai multe retractări înainte de 2017

Denumirea revistei

PLoS One

The Journal of Biological Chemistry

Immunopharmacology and Immunotoxicology

Molecular Biology Reports

Proceedings of the National Academy of Sciences of the United States of America

The Journal of Neuroscience

Tumor Biology

Archives of Biological Sciences

Diagnostic Pathology

European Journal of Medical Research

Nature

Asian Pacific Journal of Tropical Medicine

BioMed Research International

Molecular Cell

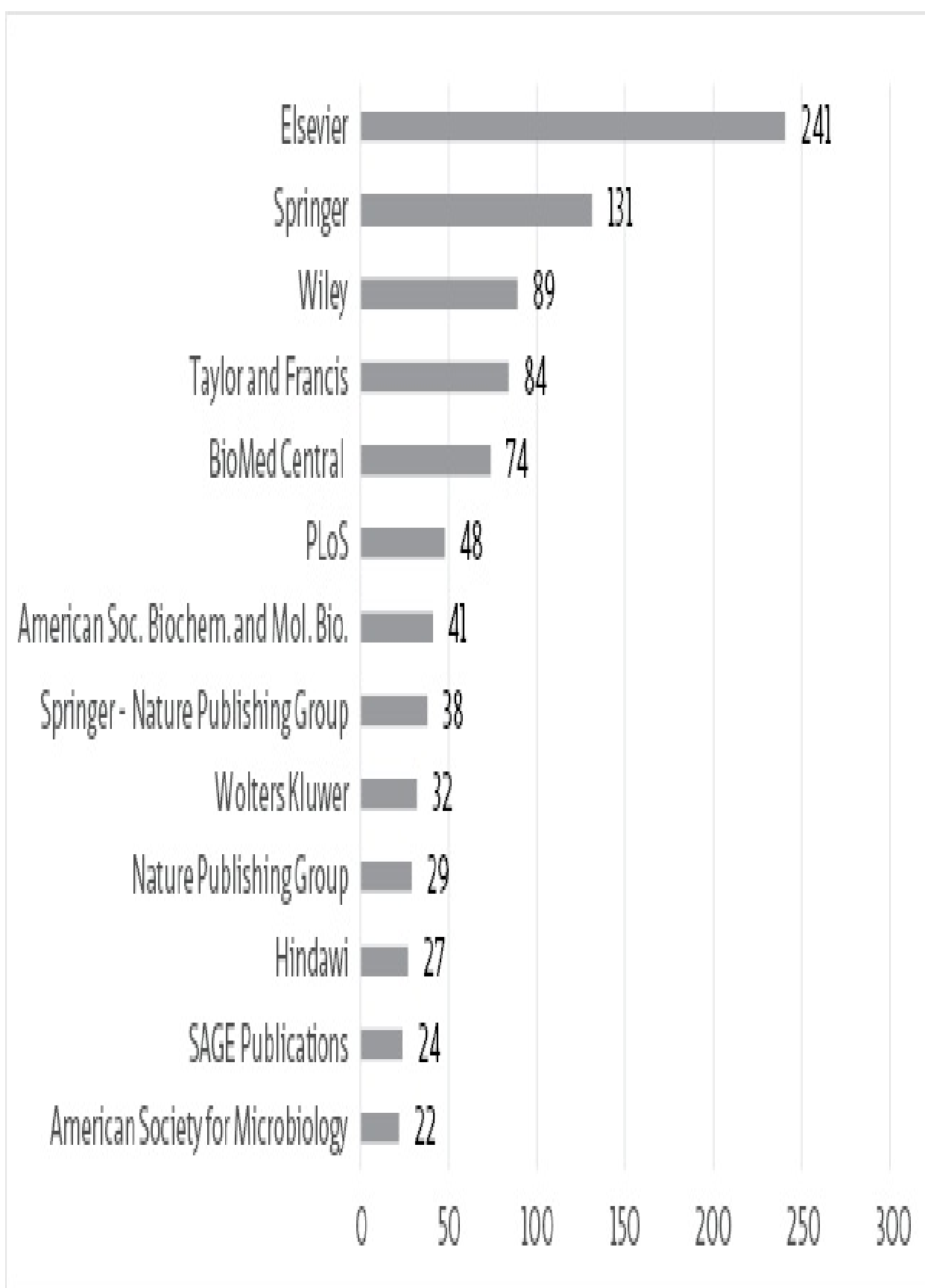
Biochemical and Biophysical Research Communications

Cell

În baza noastră de date au fost 740 de reviste. Este interesant că 2% dintre acestea sunt responsabile pentru 20% din numărul total de retractări. Tabelul 4.2 expune revistele cu cel mai mare număr de articole retractate din domeniile medical și biomedical. Deși PLoS One și The Journal of Biomedical Chemistry sunt revistele cu cele mai multe articole retractate în anii pe care i-am avut în vedere, cazul Tumor Biology merită explorat mai detaliat. În 2017, Tumor Biology a retractat peste 107 articole din cauza falselor evaluări inter pares — autorii au falsificat evaluarea fie inventând un expert, fie având semnătura unui expert, dar scriind evaluarea ei înșiși (McCook, 2017; Tumor Biology-Retractation Watch, nedatat). Volumul retractărilor din revista științifică a determinat excluderea acesteia din Web of Science și nu mai este indexată în baza de date. Numărul de articole retractate numai din acea revistă din cauza evaluărilor inter pares false indică o neglijare sistematică a standardelor etice de care ține seama orice revistă respectabilă. În acest caz, o revistă importantă ca Tumor Biology și-a neglijat complet responsabilitatea și a permis ca peste 100 de articole să fie publicate pe baza unor evaluări inter pares frauduloase.

Principalele edituri

Figura 4.4. Editurile cu cel mai mare număr de articole retractate



Sursă: creată de autori

Baza noastră de date conține articole publicate de 136 de edituri (vezi Figura 4.4). Dintre acestea, 13 (sau 10%) sunt responsabile pentru mai mult de 69% dintre retractări. Figura 4.4 indică principalele edituri, printre care Elsevier, Springer, Wiley și Taylor & Francis. Deși Nature a fuzionat cu Springer, în baza noastră de date sunt prezentate separat, fuzionarea neavând loc până la realizarea analizei. Nu e surprinzător că aceste edituri sunt în capul listei, dat fiind că sunt specializate în domeniile medical și biomedical. În ciuda faptului că Elsevier este prima dintre editurile care retractează articole din aceste domenii științifice, trebuie menționat și că Elsevier este cea mai mare editură din sfera medicală și biomedicală, având peste 2 000 de reviste dedicate acestor discipline. Springer Nature are 1 446 de reviste, Wiley, 822 și Taylor & Francis, 812 (date obținute prin Scopus).

Citări

Citările sunt considerate etalonul de aur al impactului științific. Ocuparea catedrelor, promovările și finanțările se fac în funcție de numărul de articole și de numărul de citări pe care le primesc, pe lângă prestigiul revistei în care apar. Așadar, pentru orice cercetător, citările reprezintă de departe cel mai important criteriu, dat fiind că determină soarta carierei și reputației sale. Este astfel de înțeles că presiunea de a publica este mare și cercetătorii sunt la curent cu numărul de citări pe care le primesc articolele lor. Oricine a auzit aforismul „Publică sau mori” (Publish or Perish) (atribuit în lumea științifică lui Logan Wilson (Garfield, 1996)) știe că cercetătorii se află sub o presiune enormă de a publica (Neill, 2008; Nygaard, 2017). Această presiune a fost considerată responsabilă și pentru creșterea frecvenței plagiatelor și a altor abateri din domeniul științific (Nygaard, 2017; Rawat și Meena, 2014). Baza noastră de date ne-a permis să folosim Scopus pentru a identifica citările pre

și post-retractare, întrucât aveam datele publicării inițiale a articolului și ale notificării de retractare. Deși citările pre-retractare sunt acceptabile, citările post-retractare sunt îngrijorătoare, întrucât indică faptul că un articol exclus din literatură încă e citat, adică primește recunoaștere științifică după ce a fost etichetat drept fraudulos. Tabelul 4.3 ilustrează citările pre și post-retractare în funcție de anul publicării. Datele citărilor din acest capitol au fost colectate în februarie 2019. Trebuie menționat faptul că, potrivit normelor de citare, chiar dacă un articol menționează un articol retractat într-un mod negativ, tot trebuie să îl citeze, ceea ce înseamnă că articolul fraudulos continuă să aibă un impact științific în ciuda retractării sale.

Tabelul 4.3. Citări pre și post-retractare în funcție de anul publicării

Anul publicării	Citări pre și post-retractare	Citări înainte de retractare	Citări c
2010	6 819	4 587	2 232
2011	6 104	3 650	2 454
2012	4 063	2 266	1 797
2013	2 858	1 570	1 288
2014	2 145	986	1 159
Toți anii	21 989	13 059	8 930

Audiență

Este relativ nouă capacitatea de a afla de câte ori un articol este citit. „Audiență“ face parte dintr-o categorie de evaluări denumite „altmetrics“ (alternative metrics — evaluări alternative) care au scopul de a evalua modul în care este folosită literatura științifică, pe lângă citări. Audiența (numărul de accesări) este disponibilă pe mai multe platforme, printre care Mendeley, o platformă-hibrid care funcționează ca manager de referințe și în același timp ca rețea de socializare în care oamenii de știință pot salva și recomanda articole. Se poate face legătura între audiența înregistrată pe Mendeley și citări. În primul rând, utilizatorii Mendeley sunt în cea mai mare parte oameni de știință și studenți, așa că această rețea este mai degrabă științifică, și nu o rețea de socializare deschisă, deși oricine se poate înscrie. Mai multe studii (Costas, Zahedi și Wouters, 2015; Haustein, 2016; Li și Thelwall, 2012) au găsit corelații semnificative, de intensitate medie, între audiență și numărul de citări, ceea ce indică faptul că există o suprapunere parțială între persoanele care citesc și cele care citează. Corelațiile sunt și mai puternice în domeniile medicale (Thelwall, 2016).

Tabelul 4.4. Numărul de accesări ale articolelor retractate din rețeaua Mendeley

Anul publicării	Număr de publicări	Număr de cititori	Numărul mediu de citit
2010	263	7 790	29,6
2011	260	8 896	34,2
2012	269	8 226	30,6

2013	242	12 414	51,3
2014	260	4 814	18,5
Toți anii	1 294	42 140	32,6

Din setul nostru de date, 91% dintre articolele retractate (1 178) au fost citite pe Mendeley. Am descoperit și că 1 091 (84%) dintre toate articolele retractate din setul de date aveau și citări Scopus, și accesări Mendeley. Tabelul 4.4 prezintă numărul total și numărul mediu de accesări în funcție de anul publicării. În total, numărul de cititori este destul de mare, dar iese în evidență anul 2013, unde vedem o creștere semnificativă a numărului de accesări. La bază au stat două articole publicate în acel an în Nature (vezi cazul lui Haruko Obokata) care au fost retractate din cauza falsificării și a abaterilor și care au provocat o agitație destul de mare în lumea științifică. Aceste două articole au fost cele mai citite, cu 4 130, respectiv 2 066 de accesări și au fost citate de 166, respectiv de 85 de ori. Un articol a fost citat de 78 de ori, iar celălalt, de 40 de ori după ce au fost retractate, în ciuda scandalului care a dus la sinuciderea unuia dintre autori.

Social media

Pentru a evalua impactul social al articolelor retractate, am folosit două platforme care înregistrează menționările din social media: Altmetric.com și PlumX. Când un articol este publicat primește un număr unic (DOI) care poate fi căutat și descoperit în motoarele de căutare. Și PlumX, și Altmetric.com sunt platforme bine cunoscute de înregistrare a menționărilor în social media a articolelor științifice prin numărul unic DOI (Distinct Object Identifier) și pot indica menționările pe canale social media ca Twitter, Facebook, Wikipedia și altele. Altmetric.com și PlumX înregistrează menționările din social media în modalități ușor diferite. De exemplu, când scanează bloguri și știri, Altmetric.com caută citări pe Wikipedia doar în engleză, în timp ce PlumX caută pe Wikipedia și în suedeză, sau, când numără menționările pe Facebook, Altmetric.com scanează postările de pe paginile publice, iar PlumX raportează like-uri, share-uri și comentarii combinate. O altă cercetare a descoperit că menționările pe Twitter ale

articolelor științifice sunt în general cele mai semnificative în comparație cu alte platforme de social media (Thelwall, Haustein, Larivière și Sugimoto, 2013). Cu toate acestea, identificarea menționărilor în toate platformele social media este considerată una dintre cele mai bune modalități de a evalua reacțiile publicului la literatura științifică. Tabelul 4.5 expune numărul de afișări ale articolelor retractate pe social media și pe site-urile de știri. După cum reiese din tabel, Altmetric.com a reușit să identifice 456 de articole care au fost menționate pe Twitter, în timp ce PlumX a identificat 305. Este clar că Twitter este de departe canalul dominant de social media unde articolele sunt distribuite și comentate, cu peste 7 000 de menționări identificate de Altmetric.com și aproape 6 000 găsite de PlumX. Facebook este a doua cea mai populară platformă unde aceste articole retractate au fost distribuite, discutate și comentate, având peste 10 000 de apariții conform PlumX. În plus, dat fiind că noi examinăm articolele medicale și biomedicale retractate, nu e surprinzător că și site-urile de știri au vorbit despre o parte dintre acestea. Altmetric.com a reușit să identifice 81 de articole retractate în media de știri. Aceste numere sunt în concordanță cu cele găsite de alte cercetări care au arătat că majoritatea oamenilor din SUA consumă informații medicale și științifice din social media (Greenwood et al., 2016; NW, Washington și Inquiries, 2018). Cel mai mare număr de menționări în social media și în media de știri l-a avut articolul din 2013 „Primary Prevention of Cardiovascular Disease with a Mediterranean Diet“ (Prevenție primară a tulburărilor cardiovasculare printr-o dietă mediteraneeană), publicat în New England Journal of Medicine. Acest articol a fost retractat din cauza designului și a faptului că un număr semnificativ de persoane din grupul de control s-au retras, ceea ce a afectat rezultatele. Deși același articol a fost republicat mai târziu cu o nouă analiză, efectul retractării continuă și astăzi să provoace controverse. Până acum (august 2019), acest articol retractat a fost obiectul a 195 de articole de știri, a avut peste 47 000 de share-uri, like-uri și comentarii pe Facebook și peste 1 600 de tweet-uri și retweet-uri și dezbateră continuă. Comentariile din social media pentru acest articol pot fi consultate la <https://plu.mx/judit/a/--LPrDZbZGFIXEcIgDOzjbhPj1—P4QGWWC7IKS-Um0/>.

Tabelul 4.5. Principalele date obținute prin Altmetric.com și PlumX

	Acoperire Altmetric	Suma menționărilor Altmetric	Media menționărilor Altmetric	Articole cu Altmetric
Twitter	456	7 094	15,6	2 305
Facebook	104	675	6,5	272
Știri	81	468	5,8	97
Bloguri	68	1 210	17,8	65
Wikipedia	39	49	1,3	3

Cele mai influente articole retractate

În această secțiune discutăm articolele din baza noastră de date care au fost citate de cele mai multe ori (pre și post-retractare), cele care au fost cele mai citite și cele care au avut cele mai multe tweet-uri conform Altmetric.com și PlumX.

Cel mai citat articol retractat înainte și după retractare

Estruch, R., Ros, E., Salas-Salvadó, J., Covas, M.I., Corella, D., Arós, F., ... și Lamuela-Raventos, R.M. (2013). Primary prevention of cardiovascular disease with a Mediterranean diet. *New England Journal of Medicine*, 368(14), 1279–1290.

Acest articol a fost considerat revoluționar în domeniul cercetării cardiovasculare, indicând beneficiile unei diete bazate pe uleiul de măsline în cazul tulburărilor cardiovasculare. Însă deficiențe la nivelul designului de cercetare și al descoperirilor raportate i-au făcut pe autori să scrie: „Din cauza neregularităților determinate de procedurile de randomizare, dorim să retractăm articolul” (vezi notificarea de retractare). Deși datele au fost reanalizate și articolul a fost mai târziu republicat, dezbateră în jurul beneficiilor reale ale dietei mediteraneene continuă până în ziua de azi.

Cel mai citat articol înainte de retractare

Dunoyer, P., Schott, G., Himber, C., Meyer, D., Takeda, A., Carrington, J. C și Voinnet, O. (2010). Small RNA Duplexes function as mobile silencing signals between plant cells. *Science*, 1185880.

Acest articol a fost considerat revoluționar datorită afirmației că ARN-ul care se găsește în plante poate funcționa ca apărare împotriva atacurilor virale. Articolul a fost retractat la șase ani după publicare din cauza duplicării imaginilor și a manipulărilor pe care editorii le-au considerat intenționate și nedeterminate de o eroare. Este interesant că toți autorii, cu excepția autorului principal, au fost de acord să retracteze articolul (Berg, 2016). Întrucât intervalul dintre publicare și retractare a fost de șase ani, articolul a fost citat de multe ori pre-retractare.

Cel mai citat articol post-retractare

Séralini, G.E., Clair, E., Mesnage, R., Gress, S., Defarge, N., Malatesta, M., ... și De Vendômois, J.S. (2012). Long term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize. *Food and Chemical Toxicology*, 50(11), 4221–4231.

Acest articol a fost primul care a arătat că alimentele bazate pe OMG-uri (organisme modificate genetic) sunt toxice pentru oameni și reprezintă sursa mai multor tipuri de cancer. Studiul a fost atacat de oameni de știință care au indicat mai multe deficite ale designului și metodelor de analiză și articolul a fost retractat în același an. Cu toate acestea, studiul a determinat serioase dezbateri politice provocate de activiști anti-OMG-uri care susțineau că retractarea articolului nu era validă și că a fost determinată de politică, nu de știință. Dezbaterea continuă până în ziua de azi și sunt mai multe articole

publicate recent de media de presă despre chestiunea siguranței OMG-urilor. Aceste știri și mai multe informații despre dezbateri pot fi consultate la <https://plu.mx/judit/a/-BV5QquinvXU1WNq7GKRdJQgWyZMb1bGDBtTnZaZRVU/>.

Articolul cu cel mai mare număr de cititori și de retweet-uri (Altmetric.com și PlumX)

Obokata, H., Wakayama, T., Sasai, Y., Kojima, K., Vacanti, M.P., Niwa, H., ... și Vacanti, C.A. (2014). Stimulus-triggered fate conversion of somatic cells into pluripotency. *Nature*, 505(7485), 641.

Acest articol s-a aflat în centrul unui scandal care a dus la sinuciderea unuia dintre autori. Un grup de cercetători japonezi au susținut că au descoperit o metodă de a crea celule stem pluripotente, capabile să se transforme în aproape orice alt tip de celulă prin utilizarea de stimuli fizici. După cum e de imaginat, efectul unei astfel de descoperiri, dacă ar fi fost adevărată, ar fi însemnat schimbarea întregii vieți omenești, dat fiind că ar fi reprezentat un tratament potențial pentru orice boală cunoscută. Însă acest studiu a fost rezultatul unor experimente și date falsificate. Articolul a fost retras în același an, dar a continuat să fie citit și distribuit online de peste 12 000 de ori și a fost menționat pe social media de peste 2 400 de ori. O perspectivă exhaustivă asupra acestui articol și asupra mențiunilor din social media și din știrile publicate de presa de știri pot fi găsite la https://plu.mx/judit/a/-XiygtWNXexZCjaHt_l6yDTcT9zIKiNbPqUL4-3V-Tw/.

Articole retractate foarte influente — câteva exemple recente

Probleme de analiză și de design ale cercetărilor

Séralini, G.E., Clair, E., Mesnage, R., Gress, S., Defarge, N., Malatesta, M., ... și De Vendômois, J.S. (2012). Long term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize. *Food and Chemical Toxicology*, 50(11), 4221–4231.

În 2012, un grup de oameni de știință francezi au publicat un articol referitor la nocivitatea porumbului modificat genetic. Concluziile articolului sugerau că OMG-urile ar putea cauza toate tipurile de cancer, concluzii susținute de doi ani de experimente pe șoareci. Articolul intitulat „Long term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize” („Toxicitatea pe termen lung a unui erbicid sistemic și a porumbului modificat genetic rezistent la erbicidele sistemice”) a provocat o furtună în mass-media, avertizând că alimentele care conțin OMG-uri trebuie evitate din cauza riscurilor semnificative asupra sănătății, mesaj care a ajuns rapid să fie distribuit de mii de ori pe social media și a dus la discuții aprinse. Articolul a fost retractat după ce o serie de scrisori către editori au arătat deficiența designului de cercetare, a analizei datelor și a interpretării rezultatelor, indicând astfel concluziile ca nefiind valide și susținând că cercetarea nu ar trebui folosită ca bază a politicilor de sănătate și a recomandărilor de sănătate publică. Însă acest articol și rezultatele pe care le-a raportat sunt în continuare foarte prezente în literatura științifică, în social media și la știri. În ciuda retractării din 2012, datele noastre indică citări frecvente care apar în anul 2018 (vezi Tabelul 4.6, rândul trei). Conform măsurătorilor Scopus, acest articol este în percentila 99 a numărului de citări în domeniul din care face parte, având în total 181 de citări identificate de Scopus până foarte recent. Atenția pe care studiul și rezultatele raportate au primit-o din partea social media și a știrilor este mult mai mare. PlumX a identificat aproape 8 000 de like-uri, share-uri și comentarii pe Facebook, peste 1 100 de menționări pe

Twitter și peste zece menționări la știri și pe bloguri. Am descoperit că articolul este menționat, în ciuda neajunsurilor sale, în contextele recomandărilor medicale și nutriționale (de exemplu, vezi Templeton, 2019; Ten Scientific Studies Prove that Genetically Modified Food Can Be Harmful To Human Health, 2018 (Zece studii științifice dovedesc că alimentele modificate genetic pot fi periculoase pentru sănătatea oamenilor). Trebuie să menționăm că acest articol retractat a fost publicat în aceeași formă (fără o altă rundă de evaluări inter pares) în Environmental Sciences Europe în 2014. Articolul republicat a avut parte de o atenție considerabilă din partea instituțiilor de presă și social media, de numeroși cititori pe Mendeley și de multe citări pe Scopus.

Abateri și raportări frauduloase

Obokata, H., Wakayama, T., Sasai, Y., Kojima, K., Vacanti, M.P., Niwa, H., ... și Vacanti, C.A. (2014). Stimulus-triggered fate conversion of somatic cells into pluripotency. *Nature*, 505(7485), 641.

În 2014, un grup de oameni de știință japonezi au publicat două studii pe aceeași temă și care au ieșit în evidență, despre o descoperire de pionierat în cercetarea celulelor stem, articole apărute în revista *Nature*. Principala descoperire, conform autorilor, consta în faptul că reușiseră să arate că, prin stimularea fizică, celulele adulte pot fi transformate în celule stem pluripotente care pot produce aproape orice tip de celulă. Acronimul STAP (stimulus-triggered acquisition of pluripotency, pluripotențare provocată prin aplicare de stimuli) folosit de acest grup a devenit celebru instantaneu. Semnificația acestui studiu stă în faptul că, până a fi publicat, oamenii de știință credeau că asta nu se putea realiza decât prin manipulare genetică. Dacă cercetarea era validă, consecințele ar fi reprezentat un pas uriaș spre eliminarea bolilor moștenite genetic, de exemplu. Însă la cinci luni după publicarea cercetării și în urma unei investigații întreprinse de instituția la care lucrau autorii, cercetarea a fost retractată din cauza abaterilor. Oamenii

de știință implicați au fost acuzați de slaba calitate a gestionării și înregistrării datelor, de manipularea imaginilor și figurilor și nu numai. Dată fiind amploarea acestei așa-numite descoperiri, retractarea a primit extrem de multă atenție din partea oamenilor de știință și a publicului (vezi ultimul rând din Tabelul 4.6), până în punctul în care unul dintre autori s-a sinucis în 2014, un altul a suferit un atac cerebral și un al treilea a fost internat din cauza depresiei (Goodyear, 2016). În ciuda faptului că acest articol a fost retractat relativ repede, atenția pe care i-au acordat-o social media s-a menținut și după ani întregi. Există peste 2 000 de share-uri pe Twitter și numeroase share-uri pe Facebook, dintre care majoritatea trimit la articol fără să menționeze scandalul din spatele acestuia sau faptul că a fost retras cu ani înainte. Acestea fiind spuse, unii susțin că tocmai social media au jucat un rol important în expunerea articolului (Diaz, 2018). Impactul în lumea științifică pe care l-a avut cercetarea, cu 167 de citări identificate de Scopus, este în cea mai mare parte negativ. Multe articole discută aspecte ale abaterilor științifice, probleme de reproductibilitate și principiile etice din cercetarea genetică. Autoarea principală, alături de coautori ai primelor articole, a publicat un alt articol în *Nature* în același an: „Bidirectional Developmental Potential in Reprogrammed Cells with Acquired Pluripotency” (Potențial de dezvoltare bidirecțional în celulele reprogramate cu pluripotență dobândită), care a fost retractat din motive similare în 2014. Ea a creat și un website după ce cele două articole au fost retractate (STAP HOPE PAGE, fără dată) încercând să convingă publicul de metodele ei. Website-ul nu a mai fost actualizat din 2016. Autoarea a publicat și o carte în Japonia în 2016 despre scandalul legat de celulele STAP (Knoepfler, 2016).

Probleme de reproductibilitate

Yi, P., Park, J.S. și Melton, D.A. (2013). Betatrophin: A Hormone that Controls Pancreatic β Cell Proliferation. *Cell* 53(4), 747–757).

În 2013, un grup de cercetători de la Harvard au publicat un articol care a

atras atenția și care sugera că au descoperit un nou hormon care putea crește numărul de celule beta, ceea ar fi putut vindeca diabetul de tip 1. Articolul a fost publicat într-una dintre cele mai prestigioase reviste științifice, Cell, și a fost retractat în 2017 din cauza faptului că nimeni, nici măcar autorii, nu a reușit să reproducă rezultatele experimentelor. Deși în procesul de evaluare și de retractare nu au apărut chestiuni de etică, problema reproductibilității trebuie luată în seamă. Progresul științific depinde de capacitatea de a reproduce rezultatele experimentale și, astfel, de a descoperi moduri de a avansa pe baza acestora către următorul nivel științific. Acest articol este un bun exemplu de pionierat științific care nu a dus la nimic din cauza problemelor de reproductibilitate și, deși nu a fost menționat prea des la știri sau în social media, a fost și continuă să fie citat de multe ori, având 30 de citări doar în 2018. Dacă rezultatele nu pot fi reproduse, știința e suspendată. Conform unei examinări realizate de Nature (Baker, 2016), 70% dintre cercetători au încercat fără succes să reproducă experimentele altcuiva și peste 50% nu au reușit să-și reproducă propriile experimente. Aceasta este o provocare enormă în privința nivelului general de încredere în procesul științific. Examinarea a descoperit că principalul motiv pentru care aceste experimente nu pot fi reproduse este că oamenii de știință raportează selectiv, expunând rezultatele care le susțin ipoteza, astfel încât, de exemplu, să publice mai repede și să-și asigure finanțarea. Însă criza reproductibilității (cum e numită în articol) este amplificată pentru că revistele evită să expună public această problemă, publicând replicări negative și: „De fapt, mai mulți respondenți care au publicat replicări eșuate au spus că editorii și evaluatorii le-au cerut să atenueze comparațiile cu studiul inițial“ (p. 454).

Concluzii

În 2012, Academia Națională de Științe din SUA a publicat o evaluare a 1047 de articole din domeniul biomedical și din sfera științelor organismelor vii, indexate în PubMed (Fang et al., 2012). Acest studiu a descoperit că doar 21,3% dintre articole au fost retractate din cauza unor erori neintenționate (în

opoziție cu fraudă deliberată), în timp ce 67,4% dintre retractări au fost din cauza abaterilor, printre care fraude, plagiate și duplicări. Numindu-l „epidemia de retractări“, studiul indică faptul că numărul de retractări din aceste domenii s-a înzecit din anii de după 1975. Ceea ce, pe de o parte, este un aspect pozitiv, dat fiind că demonstrează autoreglarea procesului de publicare a cercetărilor în care articolele sunt permanent analizate după publicare și, odată descoperite a fi frauduloase, sunt eliminate din literatură. Pe de altă parte, este în continuare îngrijorător faptul că astfel de articole reușesc să scape evaluatorilor și ajung să fie publicate în reviste respectabile, cum ar fi Nature, The Lancet, Cell și altele.

În acest capitol am examinat articole retractate din domeniile medical și biomedical publicate între 2010 și 2014 și retractate din numeroase motive, însă în special din cauza fraudei și abaterilor. Manipulările datelor și imaginilor, experimente neetice și analize și concluzii care induc în eroare au dus la retractări masive din multe reviste. După cum am demonstrat, intervalul de timp dintre publicare și retractare rămâne problematic. Deși acest interval tinde să scadă, acum fiind în medie de peste un an, o diferență considerabilă în comparație cu durata de zece ani din trecut, este însă un timp suficient de îndelungat pentru care aceste articole să influențeze procesul științelor medicale și biomedicale la mai multe niveluri. În primul rând, aceste articole sunt citate în literatura științifică, ceea ce înseamnă că descoperirile lor sunt aplicate în studii. Nu trebuie să uităm că în aceste domenii studiile implică pacienți înscriși în cercetări clinice sau cărora li se administrează tratamente sau medicamente. În al doilea rând, aceste articole, și mai ales cele care se pretind revoluționare, sunt citate masiv în social media și la știri, informația nefiltrată ajungând rapid la sute de mii de oameni.

După retractare, aceste articole sunt citate și discutate public pe aceste canale, ceea ce reprezintă cel mai îngrijorător fenomen specific articolelor retractate. Cauza este probabil accesibilitatea acestui conținut gratuit, disponibil pentru toată lumea, sau versiunile pre sau post-publicare care circulă liber pe platformele științifice. Și, deși s-ar putea să nu stârnească îngrijorare faptul că publicul distribuie și discută articole retractate, este cu siguranță motiv de îngrijorare continuarea citării articolelor retractate, dat fiind că s-ar putea ca oamenii să interpreteze greșit sau nici măcar să nu știe că un articol a fost retractat. În ciuda faptului că unele citări se pot arăta negative, autorii

menționând articolul pentru a demonstra că e fraudulos, cercetările noastre anterioare au arătat că multe citări sunt pozitive și se referă la studii retractate ca fiind valide (Bar-Ilan și Halevi, 2018).

Deci, ce se poate face? Cititorii trebuie să fie atenți. Indiferent dacă sunt oameni de știință sau nespecialiști, toți cei care se bazează pe articole sau știri despre studii din domeniile medical și biomedical trebuie să le evalueze conținutul. Nespecialiștii trebuie să fie atenți la sursele de finanțare și la sponsorizări. Orice cercetător trebuie să indice sursa de finanțare sau sponsorul. Dacă o organizație sau o companie este sponsorul studiului, trebuie să evaluăm cu atenție dacă studiul a fost scris pentru a susține anumite interese economice. Oamenii de știință, care cunosc bine terminologia, designul de cercetare și metoda de colectare a datelor, trebuie să evalueze articolul sub aspectul integrității datelor și imaginilor, precum și al analizei și concluziilor.

Conform ghidului COPE (Promoting integrity in research and its publication | Committee on Publication Ethics: COPE, fără dată), articolele retractate nu trebuie eliminate, ci notate clar ca fiind retractate. Însă acest ghid nu menționează dacă articolele retractate ar trebui să fie disponibile gratuit pe website-ul editurii. Am descoperit că multe articole retractate sunt disponibile gratuit, în timp ce alte articole din același număr al revistei pot fi citite doar contra cost. Faptul că sunt gratuite și disponibile determină creșterea probabilității ca oamenii care nu au acces la conținut contra cost să le citească, să le păstreze, să le distribuie și să le promoveze pe social media.

În plus, recomandăm ca editurile să elimine toate articolele retractate de pe website-ul revistelor, astfel încât cititorii să nu aibă acces la întregul text. Faptul că articolele retractate pot fi găsite integral fără costuri agravează problema. Cercetările retractate nu ar trebui să poată fi găsite nicăieri pe site-urile revistelor. Deși e clar că editurile nu pot face nimic în privința versiunilor pre-publicate disponibile pe alte platforme, ele au capacitatea și trebuie să ia măsuri în ce privește disponibilitatea articolelor pe propriile site-uri.

La fel ca orice altă formă de fake news, trebuie să fim atenți și să facem propria investigație. O sursă nouă, amplă și care permite căutări este Retraction Watch Database. Să acționăm pe baza literaturii medicale și

biomedicale false, frauduloase, ne poate costa sănătatea.

Notă

1.Newton, „Letter from Sir Isaac Newton to Robert Hooke“ (Scrisoare de la Sir Isaac Newton către Robert Hooke)

Referințe

Almeida, R.M.V., de Albuquerque Rocha, K., Catelani, F., Fontes-Pereira, A.J., și Vasconcelos, S.M. (2015). Plagiarism allegations account for most retractions in major Latin American/Caribbean databases. *Science and Engineering Ethics*, 1–10.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature News*, 533 (7604), 452. <https://doi.org/10.1038/533452a>

Bar-Ilan, J., și Halevi, G. (2017). Temporal characteristics of retracted articles: Research in progress. *ISSI 2017–16th International Conference on Scientometrics and Informetrics, Conference Proceedings*, 650–655. Accesat la www.scopus.com/inward/record.uri?eid=2-s2.0-85036606736&partnerID=40&md5=7be4a5ef9ec4266667ff48734340dd63

Bar-Ilan, J., și Halevi, G. (2018). Temporal characteristics of retracted articles. *Scientometrics*, 116 (3), 1771–1783.

Berg, J.M. (2016). Retraction. *Science*, 354 (6309), 190–190. <https://doi.org/10.1126/science.aai9397>

Biagioli, M., Kenney, M., Martin, B., și Walsh, J.P. (2018). Academic misconduct, misrepresentation and gaming: A reassessment. *Research Policy*, 48 (2), 401–413

Bozzo, A., Bali, K., Evaniew, N., și Ghert, M. (2017). Retractions in cancer research: A systematic survey. *Research Integrity and Peer Review*, 2 (1), 5.

Broad, W.J. (14 iunie 1983). Notorious darsee case shakes assumptions about science. *The New York Times*. Accesat la www.nytimes.com/1983/06/14/science/notoriousdarsee-case-shakes-assumptions-about-science.html

Budd, J.M., Sievert, M., și Schultz, T.R. (1998). Phenomena of retraction: Reasons for retraction and citations to the publications. *Journal of the American Medical Association*, 280 (3), 296–297.
<https://doi.org/10.1001/jama.280.3.296>

Budd, J.M., Sievert, M., Schultz, T.R., și Scoville, C. (1999). Effects of article retraction on citation and practice in medicine. *Bulletin of the Medical Library Association*, 87 (4), 437–443.

Cokol, M., Ozbay, F., și Rodriguez-Esteban, R. (2008). Retraction rates are on the rise. *EMBO Reports*, 9 (1), 2–2.
<https://doi.org/10.1038/sj.embor.7401143>

Corbyn, Z. (2012). Misconduct is the main cause of life-sciences retractions. *Nature*, 490 (7418), 21.

Costas, R., Zahedi, Z., și Wouters, P. (2015). Do „altmetrics“ correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, 66 (10), 2003–2019.
<https://doi.org/10.1002/asi.23309>

Diaz, A. (3 martie 2018). Science and social media. Accesat la 22 ianuarie 2019 pe website-ul Northwest Jammin: <https://medium.com/northwest-jammin/science-and-social-media-9e33ba90e04d>

Facts about the Measles Outbreak. (2 februarie 2015). The New York Times. Accesat la www.nytimes.com/interactive/2015/02/02/us/measles-facts.html

Fanelli, D., Costas, R., Fang, F.C., Casadevall, A., și Bik, E.M. (2017). Why do scientists fabricate and falsify data? A matched-control analysis of papers containing problematic image duplications. *BioRxiv*, 126805. <https://doi.org/10.1101/126805>

Fang, F.C., și Casadevall, A. (2011). Retracted science and the retraction index. *Infection and Immunity*, 79 (10), 3855–3859.

Fang, F.C., Steen, R.G., și Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences*, 109 (42), 17028–17033.

Garfield, E. (1996). What is the primordial reference for the phrase „publish or perish“. *The Scientist*, 10 (12), 11.

Goodyear, D. (22 februarie 2016). The stem-cell scandal. Accesat la www.newyorker.com/magazine/2016/02/29/the-stem-cell-scandal

Greenwood, S., Perrin, A., și Duggan, M. (11 noiembrie 2016). Social media update 2016. Accesat la 5 martie 2017, website-ul Pew Research Center: Internet, Science & Tech: www.pewinternet.org/2016/11/11/social-media-update-2016/

Halevi, G., și Bar-Ilan, J. (2016). Post retraction citations in context. *Proceedings of the Joint Workshop on Bibliometric-Enhanced Information Retrieval and Natural Language Processing for Digital Libraries (BIRNDL)*, Newark, NJ, 23–29.

Haustein, S. (2016). Grand challenges in altmetrics: Heterogeneity, data quality and dependencies. *Scientometrics*, 108 (1), 413–423. <https://doi.org/10.1007/s11192-016-1910-9>

Inoue, Y., și Muto, K. (2016). Noncompliance with human subjects' protection requirements as a reason for retracting papers: Survey of retraction notices on medical papers published from 1981 to 2011. *Accountability in*

Research, 23 (2), 123–135. <https://doi.org/10.1080/08989621.2015.1069713>

Knoepfler, P. (13 martie 2016). Haruko Obokata (小保方 晴子) launches new website with STAP claims. Accesat la 11 iulie 2019, website-ul The Niche: <https://ipscell.com/2016/03/haruko-obokata-%e5%b0%8f%e4%bf%9d%e6%96%b9-%e6%99%b4%e5%ad%90-launches-new-website-with-stap-claims/>

Li, X., și Thelwall, M. (2012). F1000, Mendeley and traditional bibliometric indicators. Proceedings of the 17th International Conference on Science and Technology Indicators, 2, 451–551. Science-Metrix and OST Montréal, Canada.

McCook, A.A. (20 aprilie 2017). A new record: Major publisher retracting more than 100 studies from cancer journal over fake peer reviews. Accesat la 2 februarie 2019, website-ul Retraction Watch: <https://retractionwatch.com/2017/04/20/new-recordmajor-publisher-retracting-100-studies-cancer-journal-fake-peer-reviews/>

Neill, U.S. (2008). Publish or perish, but at what cost? The Journal of Clinical Investigation, 118 (7), 2368–2368.

Nicholas, D., Watkinson, A., Jamali, H.R., Herman, E., Tenopir, C., Volentine, R., ... Levine, K. (2015). Peer review: Still king in the digital age. Learned Publishing, 28 (1), 15–21.

Noyori, R., și Richmond, J.P. (2013). Ethical conduct in chemical research and publishing. Advanced Synthesis & Catalysis, 355 (1), 3–9.

NW, 1615 L. St, Washington, S. 800, și Inquiries, D. 20036 U.-419–4300 | M.-419–4372 | M. (20 martie 2018). Science-related Facebook pages draw millions of followers, but only about three-in-ten posts feature new scientific discoveries. Accesat la 27 august 2018, website-ul Pew Research Center: Internet, Science & Tech: www.pewinternet.org/2018/03/21/the-science-people-see-on-social-media/ps_2018-03-21_facebook-and-science_0-01/

Nygaard, L.P. (2017). Publishing and perishing: An academic literacies framework for investigating research productivity. Studies in Higher

Education, 42 (3), 519–532.

Promoting integrity in research and its publication | Committee on Publication Ethics: COPE. (nedatat). Accesat la 11 iulie 2019, <https://publicationethics.org/>

Rawat, S., și Meena, S. (2014). Publish or perish: Where are we heading? *Journal of Research in Medical Sciences: The Official Journal of Isfahan University of Medical Sciences*, 19 (2), 87.

Redman, B.K., și Merz, J.F. (2008). Scientific misconduct: Do the punishments fit the crime? *American Association for the Advancement of Science. Science*, 321 (5890), 775.

Retraction Watch Database. (nedatat). Accesat la 18 ianuarie 2019, <http://retractiondatabase.org/RetractionSearch.aspx?>

STAP HOPE PAGE. (nedatat). Accesat la 11 iulie 2019, website-ul STAP HOPE PAGE: <https://stap-hope-page.com/>

Steen, R.G. (2010). Retractions in the scientific literature: Is the incidence of research fraud increasing? *Journal of Medical Ethics*, 37 (4), 249–253.

Steen, R.G. (2011). Retractions in the scientific literature: Do authors deliberately commit research fraud? *Journal of Medical Ethics*, 37 (2), 113–117.

Templeton, B. (22 februarie 2019). Templeton times: Cancer-linked monsanto chemical discovered in five major orange juice brands from collective evolution. Accesat la 29 iulie 2019, website-ul Templeton Times: <http://templeton01436.blogspot.com/2019/02/cancer-linked-monsanto-chemical.html>

Ten Scientific Studies Prove that Genetically Modified Food Can Be Harmful To Human Health. (13 martie 2018). Accesat la 29 iulie 2019, website-ul Global Research: www.globalresearch.ca/ten-scientific-studies-proving-gmos-can-be-harmfulto-human-health/5377054

Thelwall, M. (2016). Interpreting correlations between citation counts and other indicators. *Scientometrics*, 108 (1), 337–347.
<https://doi.org/10.1007/s11192-016-1973-7>

Thelwall, M., Haustein, S., Larivière, V., și Sugimoto, C.R. (2013). Do Altmetrics Work? Twitter and Ten Other Social Web Services. *PLoS One*, 8 (5). <https://doi.org/10.1371/journal.pone.0064841>

Tumor Biology-Retracton Watch. (nedatat). Accesat la 2 februarie 2019, <https://retractionwatch.com/category/by-journal/tumor-biology/>

Upi. (20 mai 1982). Doctor admits filing false data and is barred from U.S. support. *The New York Times*. Accesat la www.nytimes.com/1982/05/20/us/doctor-admits-filing-false-data-and-is-bared-from-us-support.html

Williams, P., și Wager, E. (2013). Exploring why and how journal editors retract articles: Findings from a qualitative study. *Science and Engineering Ethics*, 19 (1), 1–11.

** „Tratament“, în acest context, se referă la nivelul variabilei independente a studiului, deci la modul în care este prezentată informația statistică — procent din populație sau număr de articole (cf. APA, Dicționar de Psihologie Clinică, București, Trei, 2020). (N.r.)

*** Platformă deținută de Amazon care are scopul de a permite externalizarea de către companii a sarcinilor simple și repetitive, care însă nu pot fi realizate de inteligența artificială. (N.t.)

**** Subsecvență de n elemente construită pornind de la o secvență dată, în cazul acesta un eșantion de text sau vorbire. Elementele pot fi foneme, silabe, litere, cuvinte. (N.r.)

**PARTEA a II-a. Procese cognitive implicate în acceptarea, distribuirea
și corectarea dezinformării**

Capitolul 5. Când știrile (false) par adevărate

Intuirea adevărului și acceptarea și corectarea dezinformării

Norbert Schwarz și Madeline Jalbert

O analiză a 2,8 milioane de știri distribuite pe Twitter a descoperit că 59% dintre acestea au fost distribuite fără să fie citite (Gabiellkov, Ramachandran, Chaintreau și Legout, 2016). Aparent, șase din zece cititori au găsit titlul îndeajuns de interesant pentru a distribui știrea fără să o citească. În capitolul de față analizăm ceea ce face ca un mesaj să „aibă aparența“ de a fi adevărat, chiar înainte de a fi evaluat în detaliu conținutul acestuia. Mai întâi discutăm procesele psihologice de bază implicate în evaluarea adevărului unui mesaj și le ilustrăm prin câteva experimente. Apoi abordăm implicațiile acestor procese în distribuirea informației în social media și în corectarea dezinformării.

Evaluarea adevărului

Deși redistribuirea materialelor fără a fi citite le poate părea multor cititori surprinzătoare și iresponsabilă, acest fapt nu este fundamental diferit de modul în care comunicăm în viața cotidiană. În conversațiile de zi cu zi acționăm pe baza presupunerii tacite că emițătorul este un comunicator cooperant ale cărei contribuții sunt relevante pentru conversație și sunt adevărate, informative și clare (Grice, 1975; Sperber și Wilson, 1986). În afară de cazurile în care avem motive să ne îndoim că emițătorul respectă aceste reguli tacite ale conduitei conversaționale, acceptăm conținutul discursului fără multe întrebări și îl tratăm ca parte a terenului comun al conversației. Aceste procese conversaționale contribuie la numeroase erori ale judecății omenesti (pentru o analiză, vezi Schwartz, 1994, 1996). Unele cercetări sugerează chiar că înțelegerea unei afirmații presupune acceptarea cel puțin temporară a adevărului ei (Gilbert, 1991) înainte de a putea fi comparată cu dovezile relevante.

Deși este posibilă suspendarea încrederii (Hasson, Simmons și Todorov, 2005; Schul, Mayo și Burnstein, 2008), aceasta presupune ca mesajul să fie neverosimil sau ca receptorul să fie neîncrezător în momentul în care primește mesajul. Astfel, în general este favorizată acceptarea, și nu respingerea informației dacă nu există indicatori clari care să pună la îndoială intențiile celui care emite mesajul. Pentru a depăși acceptarea implicită a informației avem nevoie de motivație și de resurse cognitive, pe care este mai probabil să le investim dacă subiectul este important pentru noi și dacă există puține solicitări și distrageri exterioare. În absența acestor condiții, este probabil ca informația să fie acceptată — și uneori retransmisă — fără a fi evaluată prea atent.

Tabelul 5.1. Criteriile de adevăr

Criteriu
Compatibilitate: este compatibilă această informație cu celelalte lucruri pe care le știi?
Coerență: are coerență internă?
Credibilitate: provine dintr-o sursă credibilă?

Consens: alți oameni cred?

Dovezi: există fapte care susțin aceste informații?

Când oamenii chiar evaluează dacă este probabil ca informația să fie adevărată, în general iau în calcul o parte (rar toate) criteriile prezentate în Tabelul 5.1 (Schwarz, 2015). Este afirmația compatibilă cu celelalte lucruri pe care le știu? Are consecvență și coerență internă? Provine de la o sursă de încredere? Alți oameni o cred? Există dovezi semnificative care o susțin? Fiecare dintre aceste criterii este rațional și determină probabilitatea ca un mesaj să fie considerat adevărat. Aceste criterii pot fi evaluate prin raportarea la cunoștințele relevante, proces relativ lent și care presupune efort, uneori fiind necesară o amplă căutare de informații. Aceleași criterii pot fi evaluate și prin raportarea la propria reacție intuitivă, mai rapidă și realizată cu mai puține eforturi. Dacă reacția intuitivă inițială sugerează că s-ar putea să fie ceva în neregulă, este mai probabil ca oamenii să apeleze la o analiză mai complexă, în cazul în care timpul și circumstanțele o permit. Asta face ca reacția intuitivă inițială să fie prima evaluare a adevărului și va determina dacă persoana se va raporta în continuare la mesajul primit cu o atitudine critică sau dacă doar o va aproba tacit. Aceste presupuneri sunt compatibile cu o listă lungă de cercetări de psihologie socială (e.g. Petty și Cacioppo, 1986) și cognitivă (e.g. Kahneman, 2011; Stanovich, 1999), unde strategia care se desfășoară mai lent și presupune efort este numită adesea „analitică”, „sistematică” sau procesare prin „sistemul 2”, iar modalitatea de evaluare rapidă și intuitivă este denumită strategie „intuitivă”, „euristică” sau procesare prin „sistemul 1”.

Esențială pentru evaluarea intuitivă a adevărului este ușurința cu care mesajul poate fi procesat. De exemplu, dacă ceva este incompatibil cu alte lucruri pe care le știm sau dacă povestea este incoerentă, ne încurcăm și facem cale-ntoarsă pentru a ne asigura că am înțeles bine (Johnson-Laird, 2012; Winkielman, Huber, Kavanagh și Schwartz, 2012). Ceea ce face ca experiența subiectivă a ușurinței procesării, adesea denumită fluentă a procesării, să fie un indicator (faiibil) al posibilității ca mesajul să aibă o problemă care merită mai multă atenție. Considerații similare sunt valabile și în cazul celorlalte criterii de adevăr, după cum vom discuta mai târziu în acest capitol. În general, dificultatea procesării face ca mesajul să fie supus unei mai mari atenții, în timp ce ușurința procesării favorizează acceptarea mesajului.

Dacă ușurința sau dificultatea procesării nu ar fi determinată decât de atribute asociate semnificativ cu probabilitatea ca mesajul să fie adevărat, faptul de a ne baza pe propria experiență de procesare nu ar constitui o problemă majoră. Însă mesajele pot fi ușor sau dificil de procesat din multe motive — lectura poate fi lentă pentru că mesajul este incoerent (un criteriu relevant) sau pentru că dimensiunea literei este prea mică (ceea ce nu are legătură cu adevărul). Dat fiind că oamenii sunt mai sensibili la propriile experiențe subiective decât la sursa acelor experiențe (Schwarz, 2012), multe influențe neesențiale, care nu au legătură cu substanța mesajului, pot influența adevărul perceput. Discutăm aceste influențe neesențiale și rolul pe care îl joacă în consumul de media după ce vom trece în revistă cele cinci criterii dominante ale adevărului. După cum se va vedea, când gândurile curg lin este mai probabil ca oamenii să fie de acord fără o semnificativă analiză critică (vezi și Oyserman și Dawson, volumul de față).

„Big five“*** în evaluarea adevărului: procese analitice și intuitive**

O afirmație este mai probabil să fie acceptată ca adevărată dacă este compatibilă cu alte lucruri pe care persoana le știe, în comparație cu situația în care afirmația este în contradicție cu cunoștințele anterioare. Compatibilitatea poate fi evaluată analitic prin compararea informației cu cunoștințele deținute, ceea ce presupune motivație și timp (Petty și Cacioppo, 1986). Un indicator mai puțin solicitant este oferit de experiențele metacognitive și de reacțiile afective. Când o informație nu este consecventă convingerilor anterioare, oamenii au tendința să se încurce — durează mai mult să o citească și găsesc dificil să o proceseze (e.g. Taber și Lodge, 2006; Winkielman et al., 2012). Mai mult, informația care nu este consecventă convingerilor persoanei provoacă o reacție afectivă negativă, după cum arată cercetările asupra consecvenței cognitive (Festinger, 1957; Gawronski și Strack, 2012). Astfel, experiența de procesare și reacțiile afective ale persoanei

pot reprezenta indicatori (failibili) ai consecvenței sau inconsecvenței unei afirmații cu alte lucruri considerate adevărate.

O anumită afirmație are de asemenea o probabilitate mai mare de a fi acceptată dacă se încadrează într-un context mai larg care oferă coerență elementelor individuale, după cum se observă în cercetările asupra modelelor mentale (pentru o analiză, vezi Johnson-Laird, 2012) și în analizele asupra modalităților în care jurații ajung la concluzii (Pennington și Hastie, 1993). Coerența poate fi determinată printr-o analiză sistematică a relațiilor dintre diferitele părți ale informației declarative. Invers, poate fi evaluată în funcție de experiența de procesare: expunerile coerente sunt mai ușor de procesat decât expunerile care prezintă contradicții interne (Johnson-Laird, 2012), ceea ce face ca ușurința procesării să fie un indicator (failibil) al coerenței. Într-adevăr, oamenii se bazează pe fluența experienței când evaluează modul în care lucrurile „se leagă” (Topolinski, 2012), după cum se observă în evaluările coerenței semantice (Topolinski și Strack, 2008, 2009) și a silogismelor (Morsanyi și Handley, 2012).

Este mai probabil și ca informația să fie acceptată ca fiind adevărată când vine dintr-o sursă credibilă și de încredere. După cum arată zeci de ani de cercetări asupra persuasiunii, evaluările credibilității sursei pot fi bazate pe informația declarativă legată, de exemplu, de competența, educația, meritele sau afilierea instituțională a emițătorului și de prezența sau absența conflictului de interese (pentru analize, vezi Eagly și Chaiken, 1993; Petty și Cacioppo, 1986). Cu toate acestea, evaluările credibilității se pot baza și pe senzația de familiaritate. În viața de zi cu zi, oamenii au încredere mai mult în persoanele cunoscute decât în străini (Luhmann, 1979), de la interacțiuni personale la comerț online (Gefen, 2000). Familiaritatea rezultă din interacțiuni anterioare sau doar vederea repetată a imaginilor unei fețe este suficientă pentru a determina amplificarea percepției de onestitate, precum și acordul cu ceea ce spune persoana (Brown, Brown și Zoccoli, 2002; Weisbuch și Mackie, 2009). În mod similar, repetiția poate singură să facă un nume necunoscut să pară familiar și să-l facă pe purtătorul acestuia „să ajungă celebru peste noapte” (Jacoby, Woloshyn și Kelley, 1989), ceea ce poate, de asemenea, să ducă la creșterea competenței percepute a acestuia. Oamenii familiari sunt mai ușor de recunoscut și de amintit și numele lor devin mai ușor de pronunțat de la o întâlnire la alta. Variabilele care

influențează ușurința cu care sursa informației poate fi procesată pot așadar să amplifice credibilitatea percepută a sursei. Într-adevăr, o anumită afirmație are o probabilitate mai mare de a fi considerată adevărată dacă numele sursei este ușor de pronunțat (Newman et al., 2014).

Pentru a evalua probabilitatea ca o afirmație să fie adevărată, oamenii se gândesc și dacă ceilalți o cred — dacă mulți oameni sunt de acord, probabil că e întemeiată. Acest criteriu de consens social (Festinger, 1954) este esențial pentru multe procese de influență socială și uneori este numit principiul „dovezii sociale“ (Cialdini, 2009). După cum au indicat numeroase studii, oamenii sunt mai încrezători în propriile convingeri dacă acestea sunt împărtășite de ceilalți (Newcomb, 1943; Visser și Mirabile, 2004), au o probabilitate mai mare să susțină un mesaj dacă același lucru îl fac mulți alții (Cialdini, 2009) și au o mai mare încredere în ceea ce-și amintesc dacă mulți alții își amintesc în mod similar aceleași lucruri (Harris și Hahn, 2009; Ross, Buehler și Karr, 1998). Invers, perceperea dezacordului subminează semnificativ acceptarea mesajului, ceea ce face ca știrile despre controverse reale sau inventate să fie o strategie eficientă pentru modificarea opiniei publice (Lewandowsky, Ecker, Seifert, Schwarz și Cook, 2012; Lewandowsky, Gignac și Vaughan, 2013). Pentru a evalua amploarea consensului, oamenii pot consulta sondaje de opinie sau își pot întreba prietenii. Invers, se pot baza pe cât de „familiară“ pare convingerea — la urma urmei, toți cunoaștem mai multe convingeri populare, împărtășite de mulți, decât credințe nepopulare pe care le au puțini. Din punct de vedere empiric, informația familiară se citește mai ușor și e mai ușor de înțeles și de amintit decât informația nefamiliară, ceea ce face ca ușurința procesării să fie un indicator (faiibil) al familiarității și popularității. Astfel, modificările întâmplătoare ale ușurinței de procesare pot influența consensul perceput.

În sfârșit, încrederea se amplifică odată cu volumul dovezilor confirmatoare. Dovezile pot fi evaluate printr-o căutare exterioară, cum ar fi o analiză a literaturii științifice sau amintirea unor informații relevante; în orice caz, nivelul de încredere crește odată cu dovezile. Și reciproc, încrederea tinde să crească în funcție de cât de simplă este găsirea dovezilor — cu cât dovezile sunt mai multe, cu atât mai ușor sunt de găsit (în memorie sau în literatură). Această teorie populară este în centrul euristicii disponibilității studiate de Tversky și Kahneman (1973). Din nefericire, această euristică poate induce în

eroare. Dacă singurele informații care ne vin cu ușurință în minte apar doar pentru că au fost repetate la nesfârșit sau pentru că sar în ochi și sunt memorabile, s-ar putea să ajungem la concluzia greșită că dovezile sunt puternice. Mai mult, atenția la ceea ce ne vine în minte și atenția la ușurința cu care se întâmplă asta duc adesea la concluzii diferite. Pe de o parte, argumentele hotărâtoare care ne vin în minte ne fac să fim cu atât mai încrezători cu cât ne amintim sau generăm mai multe. Pe de altă parte, să ne bazăm pe ușurința cu care ne aducem aminte înseamnă să avem cu atât mai puțină încredere cu cât încercăm să găsim mai multe argumente, întrucât e dificil să găsim multe argumente, ceea ce sugerează că probabil nu sunt multe (Haddock, Rothman, Reber și Schwarz, 1999; pentru analize, vezi Schwartz, 1998; Schwarz și Vaughn, 2002).

Indiferent de criteriile de adevăr pe care se bazează oamenii, informațiile ușor de procesat au un avantaj în fața informațiilor dificil de procesat: par mai familiare, mai compatibile cu credințele personale, par să aibă o mai mare coerență internă, par mai larg răspândite, mai bine argumentate și par să aibă o probabilitate mai mare de a proveni de la o sursă credibilă. Aceste inferențe reflectă faptul că informațiile familiare, întâlnite frecvent, și cele coerente și compatibile cu cunoștințele anterioare sunt într-adevăr mai simplu de procesat decât celelalte tipuri de informații. Astfel, ușurința procesării oferă o euristică utilă — dar failibilă — de evaluare a gradului în care o afirmație îndeplinește principalele criterii de adevăr.

Determinarea „senzației“ de adevăr

Până acum, discuția noastră a subliniat faptul că simplitatea sau dificultatea procesării poate fi determinată și de variabile semnificative pentru criteriile esențiale de adevăr, și de influențe întâmplătoare. Ceea ce e important din două motive. Dintr-o perspectivă a cercetării, le permite cercetătorilor să manipuleze fluenta procesării în moduri independente de caracteristicile fundamentale ale mesajului și sursei acestuia. Dintr-o perspectivă practică,

subliniază faptul că afirmațiile pot să „dea senzația“ că sunt adevărate doar pentru că sunt ușor de procesat, ceea ce permite multe ocazii de manipulare. În continuare, trecem în revistă unele dintre cele mai importante variabile care influențează ușurința sau dificultatea procesării mesajului.

Repetiția

Demagogii știu de milenii întregi că adevărul poate fi creat prin repetarea frecventă a unei minciuni — după cum a spus Hitler, „[...] nicio varietate nu trebuie să modifice ținuta a ceea ce face obiectul propagandei, însă trebuie mereu [...] să reafirme același lucru“ (Hitler, 1934/1997) vol. 1, p. 172). Cercetările empirice confirmă intuiția demagogilor. Studiind zvonurile din timpul războiului, Allport și Lepkin (1945) au descoperit că cel mai bun predictor al credibilității unui zvon era numărul de expuneri ale populației la acesta. Hasher, Goldstein și Toppino (1997) au studiat această observație în laborator și le-au cerut participanților să dea scoruri în funcție de cât de adevărate credeau că sunt 60 de afirmații. Unele erau factual corecte (e.g. „Litiul este cel mai ușor metal“), altele, nu („Republica Populară Chineză a fost fondată în 1947“). Participanții au acordat scoruri în trei reprize, la intervale de două săptămâni. Între aceste sesiuni, unele afirmații au fost repetate o dată sau de două ori, iar altele, nu, existând astfel una, două sau trei expuneri. După cum era de așteptat, participanții credeau cu atât mai mult că o anumită afirmație era adevărată cu cât mai des o văzuseră, indiferent dacă era adevărată sau falsă. Numeroase studii ulterioare au confirmat puterea repetiției în cazul multor forme de conținut, de la cultură generală (e.g. Bacon, 1979) la marketing (e.g. Hawkins și Hoch, 1992) și convingeri politice (e.g. Arkes, Hackett și Boehm, 1989), intervalul de timp dintre expunere și evaluare variind de la minute (e.g. Begg și Armour, 1991) la luni (Brown și Nix, 1996). Dechêne, Stahl, Hansen și Wänke (2010) oferă o metaanaliză cuprinzătoare a acestui efect de „adevăr iluzoriu“.

Cea mai accentuată influență pe care o are repetiția este în cazul afirmațiilor în legătură cu care oamenii sunt nesiguri, dar se observă și când sunt disponibile mai multe informații despre afirmația respectivă (Fazio, Rand și

Pennycook, 2019; Unkelbach și Greifeneder, 2018). Mai rău, repetiția duce la amplificarea acordului în cazul oamenilor care chiar știu că afirmația e falsă —asta doar dacă se gândesc la ea (Fazio, Brashier, Payne și Marsh, 2015). De exemplu, repetarea afirmației „Oceanul Atlantic este cel mai mare ocean de pe Pământ“ a avut o mai mare acceptare chiar și în rândul oamenilor care știau că Pacificul e mai mare. Când afirmația repetată părea familiară, ei doar dădeau aprobator din cap, fără să o compare cu cunoștințele lor. Chiar și avertismentul că unele afirmații care le vor fi prezentate sunt false nu elimină acest efect, deși îi atenuează mărimea. Mai important, avertismentele atenuează influența repetiției doar când preced expunerea la afirmații — avertismentele oferite după ce persoanele au văzut afirmațiile nu au vreun efect sesizabil (Jalbert, Newman și Schwarz, 2019).

Repetiția duce și la amplificarea consensului social perceput, adică a percepției că o convingere este împărtășită de mulți alții. Weaver, Garcia, Schwarz și Miller (2007) le-au cerut participanților să citească opinii despre care li s-a spus că au fost preluate dintr-o discuție de grup în care o anumită opinie a fost prezentată o dată sau de trei ori. Fiecare opinie a fost atribuită unui membru al grupului. Deloc surprinzător, participanții au presupus că opinia era împărtășită de mai mulți oameni în cazul în care au citit-o ca venind de la trei membri ai grupului (72%) decât când au citit-o o singură dată (57%). Cu toate astea, faptul de a citi de trei ori o opinie a unui membru al grupului a fost aproape tot atât de influent, ducând la o estimare a consensului de 67% — aparent, o unică sursă repetitivă a părut un cor. Studii realizate mai târziu au arătat că oamenii au încredere într-un raport realizat pe baza informațiilor preluate de la martori oculari cu atât mai mult cu cât este repetat mai des, chiar și în cazul în care toate repetițiile provin de la același unic martor (Foster, Huthwaite, Yesberg, Garry și Loftus, 2012). În mod asemănător, cititorii ziarelor sunt mai încrezători în acuratețea unei relatări dacă același mesaj este prezentat în mai multe ziare, chiar dacă toate ziarele se bazează exclusiv pe același unic interviu cu aceeași unică persoană (Yousif, Aboody și Keil, 2019). Astfel de descoperiri sugerează că repetiția frecventă a acelorași formule în știrile de la televizor conferă mesajului o familiaritate care amplifică popularitatea și adevărul acestuia. Același lucru este valabil și în cazul social media, unde același mesaj continuă să apară pe măsură ce prieteni și prieteni ai prietenilor îi dau like și îl repostează, ceea ce duce la numeroase expuneri în rețea.

Dincolo de repetiție

În ciuda popularității pe care a avut-o în rândul demagogilor din trecut și pe care o are printre cei din prezent, repetiția este doar una dintre multele variabile care pot facilita ușurința procesării unei afirmații, făcând-o să pară mai comună, mai credibilă și adevărată. În continuare trecem în revistă unele dintre aceste variabile.

Reber și Schwarz (1999) au manipulat ușurința lecturii prin contrastul de culoare al fontului. În funcție de condiții, unele afirmații (e.g. „Orsono este un oraș din Chile“) erau simplu de citit datorită contrastului mare (e.g. albastru-închis pe fond alb), în timp ce altele erau greu de citit din cauza contrastului slab (e.g. albastru-deschis pe fond alb). Conform așteptărilor, aceeași afirmație avea o probabilitate mai mare de a fi considerată adevărată dacă era mai ușor de citit. În mod similar, ușurința de citire a fontului poate influența evaluările intuitive ale adevărului și măsura în care analizăm îndeaproape un mesaj. De exemplu, când sunt întrebați: „Câte animale din fiecare specie a luat Moise pe arcă?“, majoritatea oamenilor răspund „două“, chiar dacă știu că personajul biblic era Noe, nu Moise. Song și Schwarz (2008) au prezentat această întrebare despre Moise (preluată din Erickson și Mattson, 1981) cu unul din fonturile prezentate în Figura 5.1. Autorii i-au avertizat pe participanți că unele întrebări ar putea fi menite să-i inducă în eroare, caz în care ar trebui să răspundă „Nu pot să răspund“. Când a fost prezentată întrebarea despre Moise cu fontul Arial negru, ușor de citit, 88% nu au observat problema și au răspuns „Două“, în timp ce 53% au făcut asta când întrebarea a fost prezentată cu fontul Brush gri.

Alte variabile care influențează ușurința procesării au un efect similar. De exemplu, eseurile scrise de mână sunt mai atrăgătoare dacă scrisul de mână este ușor de citit (Greifeneder et al., 2010) și la fel sunt și mesajele orale dacă accentul emițătorului este ușor de înțeles (Levy-Ari și Keysar, 2010). În mod asemănător, aceeași conferință este mai puțin impresionantă dacă înregistrarea video are o slabă calitate audio, iar o conexiune slabă de telefonie în timpul unui interviu la radio cu o cercetătoare poate afecta

impresia ascultătorilor cu privire la calitatea programului ei de cercetare (Newman și Schwarz, 2018). De asemenea, oamenii consideră că o afirmație este mai probabil să fie adevărată când este prezentată într-o versiune care rimează, spre deosebire de una fără rimă, chiar dacă cele două versiuni sunt în mare măsură echivalente (McGlone și Tofighbakhsh, 2000). Chiar și o fotografie care nu are valoare de demonstrație poate determina acceptarea unei afirmații dacă face mai ușor de imaginat subiectul afirmației (pentru o analiză, vezi Newman și Zhang, volumul de față).

Pur și simplu pronunțarea fără dificultate a unui nume este suficientă pentru a-i conferi posesorului acestuia mai multă credibilitate și încredere. De exemplu, consumatorii au mai multă încredere într-un vânzător online de pe eBay dacă are un username ușor de pronunțat — e mai probabil să creadă că produsul se va ridica la înălțimea promisiunilor vânzătorului și că vânzătorul va respecta politica de returnare (Silva, Chrobot, Newman, Schwarz și Topolinski, 2017). În mod asemănător, e mai probabil ca o afirmație să fie considerată adevărată dacă numele sursei este ușor de pronunțat (Newman et al., 2014).

După cum indică această trecere selectivă în revistă, orice variabilă care poate influența ușurința procesării poate influența și judecata în privința adevărului, dat fiind că oamenii sunt foarte receptivi la experiența de procesare, dar mai puțin receptivi la sursa acestei experiențe. Când atenția este direcționată spre sursa care se întâmplă să stea la baza experienței, scade valoarea informațională a ușurinței sau dificultății resimțite și influența acesteia este atenuată sau eliminată, conform teoriei afectelor-ca-informație (pentru analize, vezi Schwarz, 2012, 2018).

Procesare analitică vs. procesare intuitivă

La fel ca în cazul altor tipuri de raționamente, este mai probabil ca oamenii să investească timpul și efortul necesar pentru procesarea atentă a informației dacă sunt suficient de motivați și dacă au timpul și ocazia de a face asta (pentru analize, vezi Greifeneder, Bless și Pham, 2011; Greifeneder și

Schwarz, 2014). Am putea spera că procesarea se va face cu atenție ori de câte ori e vorba de o chestiune importantă. Însă acest optimism nu este tocmai întemeiat. În viața cotidiană, mesajele referitoare la chestiuni pe care le considerăm importante pot ajunge la noi când suntem preocupați de altceva și nu avem oportunitatea de a ne ocupa de ele. Prin repetiție, astfel de mesaje pot deveni îndeajuns de familiare și de coerente pentru a evita o evaluare mai atentă, chiar dacă situația ne-ar permite să le acordăm mai multă atenție. După cum am menționat anterior, faptul de a le spune receptorilor că unele informații prezentate sunt false nu are o valoare protectivă decât în cazul în care avertismentul precedă prima expunere: avertismentele de mai târziu au un efect minor (Jalbert et al., 2019). În mod similar, motivația și oportunitatea de a examina critic un mesaj pot avea doar o influență limitată odată ce mesajul a fost encodat (pentru o analiză, vezi Lewandowsky et al., 2012).

Implicații pentru social media

Dinamicile evaluărilor adevărului au implicații importante pentru acceptarea și corectarea informațiilor false din lumea reală. Începând cu răspândirea televiziunii prin cablu și a radioului, cetățenii democrațiilor s-au bucurat de mai multe oportunități decât oricând de a se expune selectiv la mesaje media coerente cu perspectiva lor asupra lumii. Apariția social media este ultimul pas al acestei dezvoltări și, sub numeroase aspecte, ne putem gândi că au fost create pentru a face mesajele discutabile să pară adevărate. În primul rând, majoritatea mesajelor din social media sunt scurte, sunt scrise într-un limbaj simplu și sunt prezentate în moduri ușor de citit, ceea ce satisface multe criterii tehnice ale procesării ușoare. Aceste mesaje fluente sunt postate de prieteni, surse credibile. Conținutul pe care îl postează este de obicei compatibil cu propriile convingeri ale persoanei, dată fiind similaritatea opiniilor și valorilor în grupurile de prieteni (pentru o analiză a omofiliei de grup, vezi McPherson, Smith-Lovin și Cook, 2001). Mesajele postate primesc like-uri din partea altor prieteni, confirmându-se astfel consensul social, și sunt repostate, ceea ce duce la expuneri repetate. Cu fiecare expunere,

procesarea devine mai simplă și se amplifică percepțiile consensului social, coerenței și compatibilității. Comentariile și alte postări oferă dovezi suplimentare și amplifică și mai mult familiaritatea. În același timp, acumularea de like-uri și repostări determină mecanismul de filtrare să facă expunerea la informații opuse din ce în ce mai puțin probabilă. Site-ul Wall Street Journal „Blue Feed/Red Feed“ ilustrează modul în care mecanismul de filtrare al Facebook a dus la feed-uri (fluxuri de știri) profund diferite pentru liberali și conservatori în campania electorală din 2016 din SUA și din ce în ce mai multe cercetări analizează modul în care omofilia de opinie în cadrul rețelelor contribuie la controverse între rețele (Del Vicario et al., 2016; Gargiulo și Gandica, 2017). Îngustarea observată a dietei informaționale din social media este accentuată prin personalizarea conținutului din afara social media, unde furnizorii de internet și motoarele de căutare urmăresc interesele utilizatorilor pentru a modela informația livrată (Pariser, 2011).

Aceste procese nu numai că amplifică acceptarea afirmațiilor care par din ce în ce mai familiare și mai compatibile cu celelalte informații pe care persoana le cunoaște, ci și întrețin senzația de pricepere și de încredere. La urma urmei, o mare parte din ce vede o persoană în feed-ul său este familiar, ceea ce sugerează că știe cât de multe se pot ști despre un anumit subiect. De asemenea, informațiile nu conțin dovezi contrare, ceea ce sugerează că argumentele sunt nedispute. Aceste lucruri subliniază ceea ce Ross și Ward (1996) au descris drept „realism naiv“ — convingerea că lumea este așa cum o văd eu și oricine nu e de acord fie e dezinformat (ceea ce motivează eforturile de persuasiune), fie e rău-voitor (dacă persuasiunea eșuează). Aceste convingeri contribuie apoi la polarizare și la atribuirea reciprocă de rea-voință.

Implicații pentru corectarea dezinformării

Faptul că oamenii pot ajunge la raționamente referitoare la adevăr bazându-se mai degrabă pe strategii analitice sau intuitive reprezintă o provocare majoră

pentru campaniile de informare publică al căror scop este corectarea credințelor eronate. Cercetări ample în domeniul educației arată că noțiunile eronate ale elevilor pot fi corectate prin confruntarea lor cu informația corectă, indicându-li-se elevilor pas cu pas de ce o idee este greșită și alta este corectă, preferabil repetându-se acest proces de mai multe ori (pentru analize, vezi Vosniadou, 2008). Ceea ce funcționează cel mai bine când recipientul dorește să afle informația corectă și este suficient de motivat să fie atent, să se gândească în profunzime la subiect și să-și amintească noile informații (pentru o analiză, vezi Sinatra și Pintrich, 2003). Campaniile de informare publică urmează adesea aceste proceduri și confruntă „miturile” cu „faptele”, respectând teoriile centrate pe conținut ale învățării mesajului (McQuail, 2000; Rice și Atkin, 2001). Deși asta funcționează în sala de clasă, cu receptori motivați, cu timp suficient și cu beneficiul stimulentei, realitatea campaniilor de informare este fundamental diferită. În cazul oricărui subiect, doar unui mic segment de populație îi va păsa îndeajuns pentru a se interesa de detalii; e probabil ca majoritatea să nu vadă deloc mesajul sau să-l vadă doar în trecere și să-l proceseze superficial în timp ce face altceva. Chiar dacă oamenii își amintesc informațiile corecte când sunt testați imediat, aceasta dispare repede din memorie.

În astfel de condiții, repetarea informațiilor false pentru a le corecta poate duce mai degrabă la răspândirea de informații false în rândul unor receptori neinteresați care altfel poate nu le-ar fi întâlnit niciodată. Fără să fi procesat detaliat mesajul, acum s-ar putea să găsească afirmațiile false puțin mai familiare și mai ușor de procesat când le aud sau le văd din nou. Astfel, încercarea de a corecta credințele eronate ale câtorva oameni poate duce la pregătirea multora de a accepta aceste credințe prin expuneri repetate (pentru o analiză, vezi Schwarz, Sanna, Skurnik și Yoon, 2007). De exemplu, Skurnik, Yoon, Park și Schwarz (2005) au expus adulți tineri și maturi o dată sau de trei ori la afirmații referitoare la produse, cum ar fi: „Cartilajul de rechin este bun pentru artrită” și acestea au fost clar marcate ca „adevărate” sau „false”. Când au fost testați imediat, corectările au părut să aibă succes — toți participanții au avut o probabilitate cu atât mai scăzută de a accepta o afirmație ca fiind adevărată cu cât li s-a spus de mai multe ori că e falsă. Acesta este succesul sperat și majoritatea studiilor se opresc aici. Dar după un interval de trei zile, avertismentele repetate au o mișcare de recul și adulții maturi au o probabilitate cu atât mai mare de a considera o afirmație ca fiind

„adevărată“ cu cât li s-a spus explicit de mai multe ori că e falsă. Se presupune că receptorii nu își mai amintesc dacă afirmația a fost inițial marcată ca adevărată sau falsă, dar găsesc în continuare afirmațiile repetate mai ușor de procesat și mai familiare, ceea ce a făcut ca acestea să „pară“ adevărate.

Chiar și expunerea oamenilor doar la informații adevărate poate crește probabilitatea de a accepta o versiune falsă a acelor informații odată cu trecerea timpului. Garcia-Marques, Silva, Reber și Unkelbach (2015) le-au prezentat participanților afirmații ambigue (e.g. „crocodilii dorm cu ochii închiși“ și mai târziu le-au cerut să indice dacă sunt adevărate afirmații care erau fie identice cu cele văzute anterior, fie erau în contradicție directă cu acestea (e.g. „crocodilii dorm cu ochii deschiși“). Când participanții au făcut aceste judecăți imediat, au indicat afirmațiile identice repetate ca fiind adevărate și afirmațiile care le contraziceau ca fiind false în mai mare măsură decât afirmațiile noi, pe care nu le mai văzuseră. Însă după o săptămână, au considerat și afirmațiile identice, și cele care le contraziceau ca fiind adevărate în mai mare măsură decât afirmațiile noi. Simplu spus, dacă intervalul de timp este suficient de scurt, oamenii își amintesc exact informațiile pe care tocmai le-au văzut și le resping pe cele opuse. Însă, pe măsură ce timpul trece, detaliile se pierd și informațiile contradictorii par mai familiare decât informațiile care n-au mai fost întâlnite până atunci — da, era ceva despre crocodili și despre ochii lor, așa că probabil despre asta era vorba.

Pe măsură ce timpul trece, oamenii pot chiar să infereze credibilitatea sursei inițiale pe baza intensității convingerii. De exemplu, Fragale și Heath (2004) au expus participanții de două până la cinci ori la afirmații precum: „S-a demonstrat că ceapa folosită la izolarea recipientelor Cup-o-Noodles provoacă cancer în cazul șobolanilor“. Apoi, participanții au aflat că unele afirmații au fost preluate din National Enquirer (sursă cu credibilitate scăzută) și altele, din Consumer Reports (sursă cu credibilitate mare) și trebuiau să indice sursa probabilă a fiecărei afirmații. Cu cât mai des au auzit participanții o afirmație, cu atât era mai mare probabilitatea de a o atribui Consumer Reports, și nu National Enquirer. Pe scurt, expunerea frecventă nu determină doar amplificarea convingerii că o afirmație e adevărată, ci amplifică și convingerea că sursa este de încredere. În mod similar, eforturile bine-

intenționate ale Centers for Disease Control și Los Angeles Times de a infirma un zvon despre „bananele mâncătoare de carne“ s-a transformat în convingerea că Los Angeles Times îi avertiza pe oameni să nu mănânce acele banane periculoase, zvonul fiind astfel amplificat (Emery, 2000). Astfel de erori de atribuire a sursei cresc probabilitatea ca oamenii să le transmită informația și altora, care la rândul lor au o probabilitate mai mare de a o accepta (și de a o răspândi), dată fiind presupusa sursă credibilă (Rosnow și Fine, 1976).

Astfel de descoperiri ilustrează faptul că încercările de corectare a dezinformării pot avea un efect de recul dacă se concentrează exclusiv asupra conținutului mesajului în defavoarea efectului pe care mesajul îl are asupra experienței ulterioare de procesare a receptorilor. Chiar dacă un mesaj corector reușește să schimbe convingerea receptorilor care sunt profund interesați de subiect și care procesează mesajul cu suficientă atenție, s-ar putea să răspândească informații false în rândul multor alți oameni cărora nu le pasă de acest subiect. Din nefericire, e probabil ca aceștia din urmă să fie mult mai mulți decât primii. În acest caz, corectarea reușită a câtorva convingeri eronate se poate face cu prețul inducerii în eroare a numeroase persoane neimplicate. Pentru a evita astfel de efecte de recul, în general e mai sigură abținerea de la orice reiterare a informațiilor false și concentrarea exclusivă asupra faptelor. Cu cât faptele devin mai familiare și mai cursive, cu atât este mai probabil să fie acceptate ca adevărate și să servească drept bază pentru judecăți și decizii (Lewandowsky et al., 2012; Schwarz et al., 2007; 2016).

Din nefericire, adevărul este în general mai complicat decât poveștile amăgitoare, care adesea implică simplificări considerabile. Ceea ce face ca adevărul să fie dezavantajat, întrucât e mai dificil de procesat, de înțeles și de amintit. Este așadar importantă prezentarea informației în moduri care facilitează procesarea cursivă. Asta implică expuneri pas-cu-pas și evitarea jargonului. Ajută și să acordăm atenție influențelor din mediu asupra ușurinței procesării. Un font ușor de citit și o pronunțare ușor de înțeles, adăugarea de fotografii și repetarea aspectelor esențiale sunt toate tehnici care nu ar trebui lăsate doar în seama celor care doresc să inducă în eroare — acestea pot să-i ofere adevărului o mână de ajutor și ar trebui folosite.

În sfârșit, la nivel individual, cea mai bună apărare împotriva influenței dezinformării este scepticismul în primul moment în care informația este întâlnită (pentru o analiză, vezi Lewandowsky et al., 2012). Odată ce oamenii au procesat informațiile false, avertismentele au o influență slabă. Pe lângă avertismentele explicite, o atitudine generală de suspiciune și neîncredere face ca mesajul să fie analizat mai atent și scade nivelul de acceptare a mesajului (pentru analize, vezi Mayo, 2017; Schwarz și Lee, 2019). Avertismentele explicite, precum și suspiciunea și neîncrederea implică faptul că cel care comunică s-ar putea să nu adere la normele de cooperare conversațională (Grice, 1975), mesajul necesitând astfel mai multă atenție. Din nefericire, într-un climat de opinie publică polarizată, simplul fapt că un mesaj susține „cealaltă” tabără poate să reprezinte în sine un motiv de suspiciune și neîncredere, ceea ce afectează încercările de corectare în contexte polarizate.

Mulțumiri

Pregătirea acestui capitol a fost susținută de Linnie și Michael Katz Endowed Research Fellowship Fund printr-o bursă de cercetare acordată celui de-al doilea autor și de fonduri acordate primului autor prin USC Dornsife Mind and Society Center.

Referințe

Allport, F.H., și Lepkin, M. (1945). Wartime rumors of waste and special privilege: Why some people believe them. *Journal of Abnormal and Social Psychology*, 40, 3–36.

Arkes, H.R., Hackett, C., și Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, 2, 81–94.

Bacon, F.T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 241–252.

Begg, I., și Armour, V. (1991). Repetition and the ring of truth: Biasing comments. *Canadian Journal of Behavioural Science*, 23, 195–213.

Brown, A.S., Brown, L.A., și Zoccoli, S.L. (2002). Repetition-based credibility enhancement of unfamiliar faces. *The American Journal of Psychology*, 115, 199–2009.

Brown, A.S., și Nix, L.A. (1996). Turning lies into truths: Referential validation of falsehoods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1088–1100.

Cialdini, R.B. (2009). *Influence: Science and practice*. Boston: Pearson Education.

Dechêne, A., Stahl, C., Hansen, J., și Wänke, M. (2010). The truth about the truth: A metaanalytic review of the truth effect. *Personality and Social Psychology Review*, 14, 238–257.

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., ... Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113 (3), 554–559.

Eagly, A.H., și Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace.

Emery, D. (23 februarie 2000). The great banana scare of 2000. Accesat la 24 mai, 2002, <http://urbanlegends.about.com/library/weekly/aa022302a.htm>

Erickson, T.D., și Mattson, M.E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning & Verbal Behavior*, 20, 540–551.

Fazio, L.K., Brashier, N.M., Payne, B.K., și Marsh, E.J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General*, 144 (5), 993–1002.

Fazio, L.K., Rand, D.G., și Pennycook, G. (2019). Repetition increases perceived truth equally for plausible and implausible statements. *Psychonomic Bulletin & Review*, 26 (5), 1705–1710.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 123–146.

Festinger, L. (1957). A theory of cognitive dissonance. Evanston, IL: Row, Peterson.

Foster, J.L., Huthwaite, T., Yesberg, J.A., Garry, M., și Loftus, E. (2012). Repetition, not number of sources, increases both susceptibility to misinformation and confidence in the accuracy of eyewitnesses. *Acta Psychologica*, 139, 320–326.

Fragale, A.R., și Heath, C. (2004). Evolving information credentials: The (mis)attribution of believable facts to credible sources. *Personality and Social Psychology Bulletin*, 30, 225–236.

Gabielkov, M., Ramachandran, A., Chaintreau, A., și Legout, A. (2016). Social clicks: What and who gets read on Twitter? *ACM SIGMETRICS Performance Evaluation Review*, 44, 179–192.
<http://dx.doi.org/10.1145/2896377.2901462>

Garcia-Marques, T., Silva, R.R., Reber, R., și Unkelbach, C. (2015). Hearing a statement now and believing the opposite later. *Journal of Experimental Social Psychology*, 56, 126–129.

Gargiulo, F., și Gandica, Y. (2017). The role of homophily in the emergence of opinion controversies. *Journal of Artificial Societies and Social Simulation*, 20 (3), 8. doi: 10.18564/jasss.3448. Accesat la <http://jasss.soc.surrey.ac.uk/20/3/8.htm>

Gawronski, B., și Strack, F. (ed.). (2012). Cognitive consistency: A

fundamental principle in social cognition. New York: Guilford Press.

Gefen, D. (2000). E-commerce: The role of familiarity and trust. *Omega*, 28, 725–737.

Gilbert, D.T. (1991). How mental systems believe. *American Psychologist*, 46, 107–119.

Greifeneder, R., Alt, A., Bottenberg, K., Seele, T., Zelt, S., și Wagener, D. (2010). On writing legibly: Processing fluency systematically biases evaluations of handwritten material. *Social Psychological and Personality Science*, 1, 230–237.

Greifeneder, R., Bless, H., și Pham, M.T. (2011). When do people rely on cognitive and affective feelings in judgment? A review. *Personality and Social Psychology Review*, 15, 107–141.

Greifeneder, R., și Schwarz, N. (2014). Metacognitive processes and subjective experience. În J.W. Sherman, B. Gawronski, și Y. Trope (coord.), *Dual-process theories of the social mind* (pp. 314–327). New York, NY: Guilford Press.

Grice, H.P. (1975). Logic and conversation. În P. Cole și J.L. Morgan (coord.), *Syntax and semantics*, vol. 3: Speech acts (pp. 41–58). New York: Academic Press.

Haddock, G., Rothman, A.J., Reber, R., și Schwarz, N. (1999). Forming judgments of attitude certainty, importance, and intensity: The role of subjective experiences. *Personality and Social Psychology Bulletin*, 25, 771–782.

Harris, A.J.L., și Hahn, U. (2009). Bayesian rationality in evaluating multiple testimonies: Incorporating the role of coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1366–1372.

Hasher, L., Goldstein, D., și Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning & Verbal Behavior*, 16, 107–112.

Hasson, U., Simmons, J.P., și Todorov, A. (2005). Believe it or not: On the possibility of suspending belief. *Psychological Science*, 16, 566–571.

Hawkins, S.A., și Hoch, S.J. (1992). Low-involvement learning: Memory without evaluation. *Journal of Consumer Research*, 19, 212–225.

Jacoby, L.L., Woloshyn, V., și Kelley, C.M. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, 118, 115–125.

Jalbert, M., Newman, E.J., și Schwarz, N. (2019). Only half of what I tell you is true: How experimental procedures lead to an underestimation of the truth effect. Manuscris în curs de evaluare.

Johnson-Laird, P.N. (2012). Mental models and consistency. În B. Gawronski și F. Strack (coord.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 225–243). New York: Guilford Press.

Kahneman, D. (2012). *Gândire rapidă, gândire lentă*. București: Publica.

Levy-Ari, S., și Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, 46, 1093–1096.

Lewandowsky, S., Ecker, U.K.H., Seifert, C., Schwarz, N., și Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13, 106–131.

Lewandowsky, S., Gignac, G.E., și Vaughan, S. (2013). The pivotal role of perceived scientific consensus in acceptance of science. *Nature Climate Change*, 3, 399–404.

Luhmann, N. (1979). *Trust and power*. Chichester, Marea Britanie: Wiley.

Mayo, R. (2017). Cognition is a matter of trust: Distrust tunes cognitive processes. *European Review of Social Psychology*, 26, 283–327.

McGlone, M.S., și Tofighbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, 11, 424–428.

McPherson, M., Smith-Lovin, L., și Cook, J.M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415–444.

McQuail, D. (2000). *McQuail's mass communication theory*. Newbury Park, CA: Sage Publications.

Morsanyi, K., și Handley, S.J. (2012). Logic feels so good: I like it! Evidence for intuitive detection of logicality in syllogistic reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 596–616.

Newcomb, T.M. (1943). *Personality and social change*. New York: Holt, Rinehart, & Winston.

Newman, E.J., Sanson, M., Miller, E.K., Quigley-McBride, A., Foster, J.L., Bernstein, D.M., și Garry, M. (2014). People with easier to pronounce names promote truthiness of claims. *PLoS One*, 9 (2). doi: 10.1371/journal.pone.0088671

Newman, E.J., și Schwarz, N. (2018). Good sound, good research: How audio quality influences perceptions of the researcher and research. *Science Communication*, 40 (2), 246–257.

Newman, E.J., și Zhang, L. (2020). Truthiness: How nonprobative photos shape beliefs. În R. Greifeneder, M. Jaffé, E.J. Newman și N. Schwarz (ed.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 90–114). Londra, Marea Britanie: Routledge.

Oyserman, D., și Dawson, A. (2020). Your fake news, our fakes: Identity-based motivation shapes what we believe, share, and accept. În R. Greifeneder, M. Jaffé, E.J. Newman, și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 173–195). Londra, Marea Britanie: Routledge.

Pariser, E. (2011). *The filter bubble: How the new personalized web is*

changing what we read and how they think. New York: Penguin Books.

Pennington, N., și Hastie, R. (1993). The story model for juror decision making. În R. Hastie (coord.), *Inside the juror* (pp. 192–223). New York: Cambridge University Press.

Petty, R.E., și Cacioppo, J.T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19, 123–205.

Reber, R. și Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, 8, 338–342.

Rice, R. și Atkin, C. (coord.). (2001). *Public communication campaigns* (ed. a III-a). Newbury Park, CA: Sage Publications.

Rosnow, R.L. și Fine, G.A. (1976). *Rumor and gossip: The social psychology of hearsay*. New York: Elsevier.

Ross, L. și Ward, A. (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. În E.S. Reed, E. Turiel și T. Brown (coord.), *Values and knowledge* (pp. 103–135). Hillsdale, NJ: Lawrence Erlbaum.

Ross, M., Buehler, R., și Karr, J.W. (1998). Assessing the accuracy of conflicting autobiographical memories. *Memory and Cognition*, 26, 1233–1244.

Schul, Y., Mayo, R., și Burnstein, E. (2008). The value of distrust. *Journal of Experimental Social Psychology*, 44, 1293–1302.

Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation. *Advances in Experimental Social Psychology*, 26, 123–162.

Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods, and the logic of conversation*. Hillsdale, NJ: Erlbaum.

Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality*

and Social Psychology Review, 2, 87–99.

Schwarz, N. (2012). Feelings-as-information theory. În P.A. Van Lange, A.W. Kruglanski, și E. Higgins (coord.), *Handbook of theories of social psychology* (pp. 289–308). Thousand Oaks, CA: Sage Publications.

Schwarz, N. (2015). Metacognition. În M. Mikulincer, P.R. Shaver, E. Borgida, și J.A. Bargh (coord.), *APA handbook of personality and social psychology: Attitudes and social cognition* (pp. 203–229). Washington, DC: APA.

Schwarz, N. (2018). Of fluency, beauty, and truth: Inferences from metacognitive experiences. În J. Proust și M. Fortier (coord.), *Metacognitive diversity: An interdisciplinary approach* (pp. 25–46). New York: Oxford University Press.

Schwarz, N. și Lee, S.W.S. (2019). The smell of suspicion: How the nose curbs gullibility. În J.P. Forgas și R.F. Baumeister (coord.), *The social psychology of gullibility: Fake news, conspiracy theories, and irrational beliefs* (pp. 234–252). New York: Routledge and Psychology Press.

Schwarz, N., Newman, E. și Leach, W. (2016). Making the truth stick and the myths fade: Lessons from cognitive psychology. *Behavioral Science & Policy*, 2 (1), 85–95.

Schwarz, N., Sanna, L.J., Skurnik, I. și Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology*, 39, 127–161.

Schwarz, N., și Vaughn, L.A. (2002). The availability heuristic revisited: Ease of recall and content of recall as distinct sources of information. În T. Gilovich, D. Griffin și D. Kahneman (coord.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 103–119). Cambridge: Cambridge University Press.

Silva, R.R., Chrobot, N., Newman, E., Schwarz, N., și Topolinski, S. (2017). Make it short and easy: Username complexity determines trustworthiness

above and beyond objective reputation. *Frontiers in Psychology*, 8, 2200.

Sinatra, G.M. și Pintrich, P. (2003). The role of intentions in conceptual change learning. În G.M. Sinatra și P.R. Pintrich (coord.), *Intentional conceptual change*. Mahwah, NJ: Lawrence Erlbaum Associates.

Skurnik, I., Yoon, C., Park, D.C. și Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, 31, 713–724.

Song, H., și Schwarz, N. (2008). Fluency and the detection of misleading questions: Low processing fluency attenuates the Moses illusion. *Social Cognition*, 26, 791–799.

Sperber, D. și Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.

Stanovich, K.E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah: Erlbaum.

Taber, C.S. și Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50 (3), 755–769.

Toland, J. (1976). *Adolf Hitler*. Garden City, NY: Doubleday.

Topolinski, S. (2012). Nonpropositional consistency. În B. Gawronski și F. Strack (coord.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 112–131). New York: Guilford Press.

Topolinski, S. și Strack, F. (2008). Where there's a will: There's no intuition: The unintentional basis of semantic coherence judgments. *Journal of Memory and Language*, 58, 1032–1048.

Topolinski, S. și Strack, F. (2009). The architecture of intuition: Fluency and affect determine intuitive judgments of semantic and visual coherence and judgments of grammaticality in artificial grammar learning. *Journal of Experimental Psychology: General*, 138, 39–63.

Tversky, A. și Kahneman, D. (1973). Availability: A heuristic for judging

frequency and probability. *Cognitive Psychology*, 5, 207–232.

Unkelbach, C. și Greifeneder, R. (2018). Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology*, 79, 78–86.

Visser, P.S., și Mirabile, R.R. (2004). Attitudes in the social context: The impact of social network composition on individual-level attitude strength. *Journal of Personality and Social Psychology*, 87, 779–795.

Vosniadou, S. (coord.). (2008). *International handbook of research on conceptual change*. New York, NY: Routledge.

Weaver, K., Garcia, S.M., Schwarz, N. și Miller, D.T. (2007). Inferring the popularity of an opinion from its familiarity: A repetitive voice can sound like a chorus. *Journal of Personality and Social Psychology*, 92, 821–833.

Weisbuch, M. și Mackie, D. (2009). False fame, perceptual clarity, or persuasion? Flexible fluency attribution in spokesperson familiarity effects. *Journal of Consumer Psychology*, 19 (1), 62–72.

Winkielman, P., Huber, D.E., Kavanagh, L. și Schwarz, N. (2012). Fluency of consistency: When thoughts fit nicely and flow smoothly. În B. Gawronski și F. Strack (coord.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 89–111). New York: Guilford Press.

Yousif, S.R., Aboody, R. și Keil, F.C. (2019). The illusion of consensus: A failure to distinguish between true and false consensus. *Psychological Science*, 30 (8), 1195–1204.

Capitolul 6. Aparența de adevăr

Modul în care fotografiile neconcludente modelează convingeri

Eryn J. Newman și Lynn Zhang

„Ultima oră: zeci de mii de voturi frauduloase pro-Clinton descoperite într-un depozit din Ohio.“ La sfârșitul anului 2016, aproximativ șase milioane de oameni au fost expuși la această afirmație în social media și pe alte site-uri de știri, afirmație care a apărut alături de fotografia unui bărbat care stătea în fața unor cutii cu buletine de vot.¹ Era vorba de fake news, știrea fiind inventată de un student. Iar fotografia? O poză generică (stock photo) pe care același student a găsit-o căutând pe google „cutii cu buletine de vot“. Acest exemplu nu este unic. Fake news și dezinformarea sunt adesea însoțite de fotografii decorative care au legătură cu subiectul general, dar nu oferă nicio dovadă concludentă că titlul ar fi corect (pentru mai multe exemple, vezi Politico Staff, 2017). Dintr-o perspectivă a comunicării, utilizarea de stock photos are sens: fotografiile atrag atenția și, uneori, pot să contribuie la înțelegere și pot să amplifice probabilitatea ca oamenii să-și amintească conținutul asociat (Carney și Levin, 2002; Knobloch, Hastall, Zillmann și Callison, 2003; Waddill și McDaniel, 1992). Dar decorarea titlurilor sau afirmațiilor cu aceste fotografii neconcludente poate avea efecte mai insidioase asupra convingerilor pe care le au oamenii.

Din ce în ce mai multe cercetări arată că până și o expunere minoră la o fotografie care are legătură cu subiectul, dar care nu e concludentă, îi poate

face pe oameni să creadă că o afirmație asociată este adevărată, în ciuda faptului că fotografia nu oferă dovezi semnificative în privința veridicității afirmației, un efect al „senzației de veridicitate” (truthiness effect)² (Fenn, Newman, Pezdek și Garry, 2013; Newman, Garry, Bernstein, Kantner și Lindsay, 2012; Newman et al., 2015). Acest efect durează câteva zile și influențează mai multe tipuri de raționamente, inclusiv judecăți referitoare la cunoștințe generale, la predicții despre evenimente viitoare și judecăți cu privire la propriile amintiri episodice. Vom trece în revistă literatura referitoare la senzația de veridicitate, consemnând modurile în care fotografiile și alte tipuri de informații neconcludente pot să schimbe rapid convingerile, amintirile și judecățile oamenilor asupra propriilor cunoștințe generale. Examinăm și mecanismele care contribuie la apariția senzației de veridicitate și explorăm implicațiile pe care le au asupra fake news și a dezinformării.

Fotografii neconcludente și senzația de veridicitate

Este bine cunoscut faptul că fotografiile înșelătoare ne pot amăgi. Dacă sunt manipulate pentru a prezenta un eveniment care nu a avut loc sau dacă apar în combinație cu afirmații repetate dintr-o sursă de încredere îi pot face pe oameni să creadă și să-și amintească informații complet false (Lindsay, Hagen, Read, Wade și Garry, 2004). Dar fotografiile precum cea cu cutiile cu buletine de vot, care nu sunt decât decorative, privite doar în trecere, au fost considerate relativ inofensive (e.g. Carney și Levin, 2002). Ce dovezi avem că astfel de fotografii decorative, neconcludente, au vreun efect asupra convingerilor?

Cercetări experimentale asupra senzației de

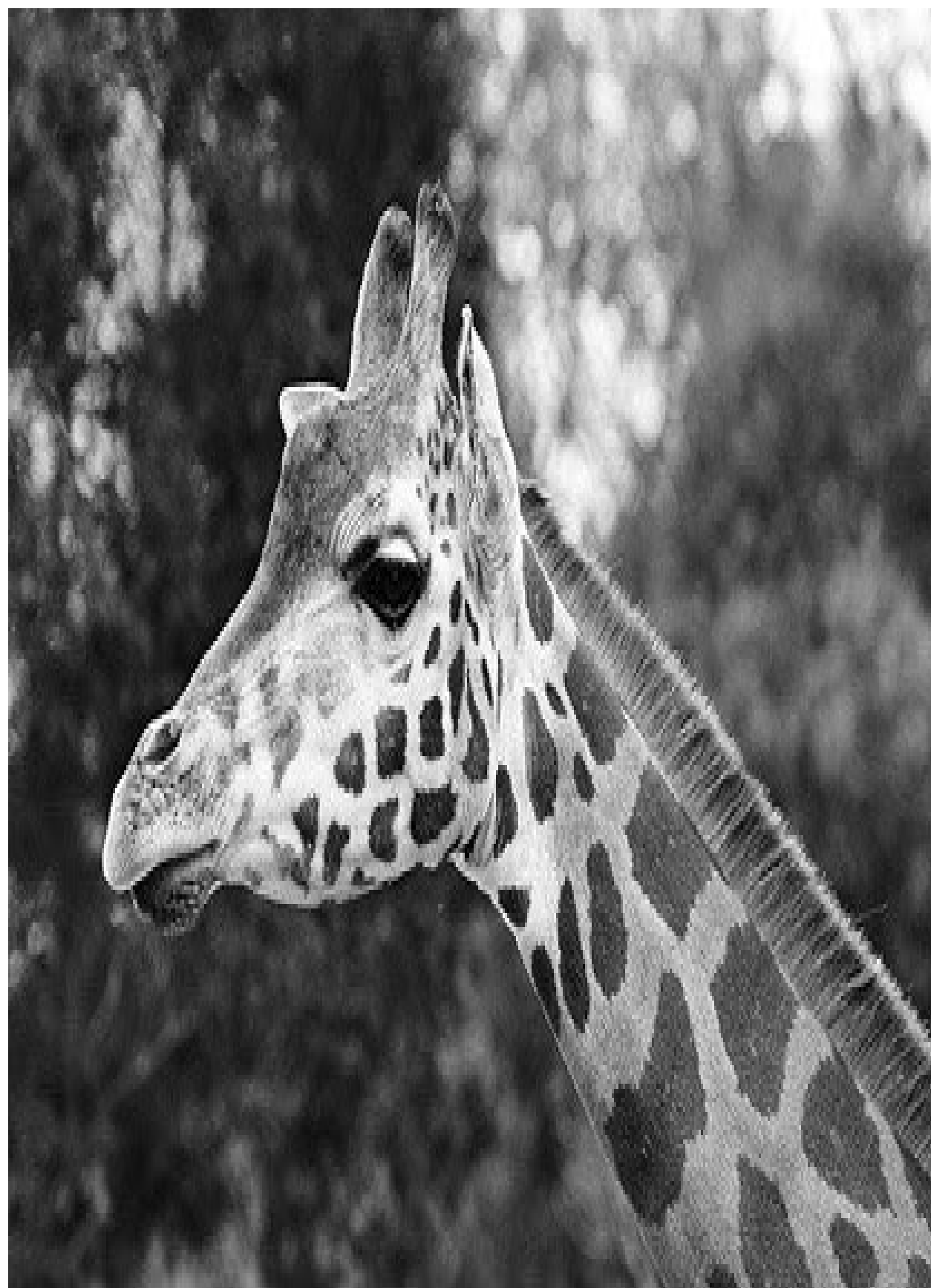
veridicitate

Într-unul dintre primele experimente care au examinat influența fotografiilor neconcludente, participanților li s-a cerut să dea un test de cunoștințe generale, unde au văzut o serie de afirmații care apăreau pe un ecran de computer (Newman et al., 2012). În acest experiment manipularea consta în faptul că jumătate dintre afirmații apăreau alături de o fotografie neconcludentă, cu un format asemănător celui din ziare sau din social media, și jumătate apăreau fără fotografii. De exemplu, participanții au văzut afirmații precum „Girafele sunt singurele mamifere care nu pot să sară” prezentate fie alături de o fotografie, ca poza care înfățișază capul unei girafe din Figura 6.1, fie fără fotografie. În ciuda faptului că fotografiile nu ofereau nicio dovadă că afirmațiile ar fi fost adevărate sau nu — poza capului unei girafe nu spune deloc dacă girafele pot sau nu să sară —, prezența unei fotografii a amplificat tendința oamenilor de a spune că afirmațiile asociate erau adevărate. Fotografiile au provocat senzația de veridicitate, o tendință de a crede afirmații prezentate alături de informații neconcludente. Într-o altă serie de experimente, publicate în același articol, Newman și colegii săi au replicat conceptual rezultatul. În aceste experimente, participanților li s-a cerut să joace un alt joc de cunoștințe generale: „Mort sau viu” (joc pe care unul dintre autori și-l amintea din vechile programe de la radio). Sarcina esențială era stabilirea faptului dacă afirmația „Această persoană este în viață” era adevărată sau falsă în cazul fiecărui nume celebru care apărea pe ecran. În jumătate dintre cazuri, numele celebrităților apăreau fără o fotografie concludentă — o fotografie cu acea persoană celebră implicată în propria profesie, dar neoferind nicio dovadă pentru adevărul afirmației „Această persoană este în viață”. De exemplu, subiecții ar fi putut vedea numele „Nick Cave” alături de o fotografie cu Nick Cave pe scenă, având în mână un microfon și cântând în fața unei mulțimi (vezi Figura 6.1). Fotografia nu oferea niciun indiciu în ce privește faptul dacă Nick Cave era în viață sau murise. Cele mai multe fotografii erau poze generice cu celebrități. Rezultatele acestui experiment au fost clare: era mai probabil ca oamenii să accepte afirmația „Această persoană este în viață” ca fiind adevărată dacă numele celebrității apărea alături de o fotografie, în comparație cu cazul în

care nu era prezentă o fotografie. Poate mai surprinzător, același tipar de rezultate a fost descoperit când unui alt grup de subiecți li s-au arătat aceleași nume de celebrități, cu aceleași fotografii, dar au evaluat afirmația opusă: „Această persoană este moartă“. Cu alte cuvinte, aceleași fotografii i-au făcut pe oameni să creadă nu numai afirmațiile că celebritățile erau „vii“, ci și afirmațiile că aceleași persoane erau „moarte“.

Figura 6.1. Sus este un exemplu de fotografie neconcludentă cu Nick Cave din experimentele „Mort sau viu“ (Newman et al., 2012); jos, un exemplu de fotografie neconcludentă din experimentul cu afirmații de cultură generală expus de Newman et al. (2012, 2015).

Sursă: fotografia cu Nick Cave: Creative Commons License attribution: Marco Annunziata. Girafa: fotografie de Fir0002/Flagstaffotos, conform GFDL v1.2 www.gnu.org/licenses/old-licenses/fdl-1.2.en.html





Acest efect al senzației de veridicitate — conform căruia fotografii neconcludente îi fac pe oameni să creadă că anumite afirmații sunt adevărate — este prezent și în alte tipuri de raționamente. În cadrul unei sarcini de predicție a pieței mărfurilor, Newman, Azad, Lindsay și Garry (2018) le-au cerut participanților să ghicească dacă afirmațiile despre diferitele mărfuri erau adevărate sau false. Participanții vedeau nume de mărfuri, ca „Benzen“, apărând pe ecran și trebuiau să spună dacă afirmația „prețul va crește“ era adevărată sau falsă. La fel ca în studiile anterioare care au folosit informații de cultură generală, afirmațiile despre mărfuri apăreau pe ecran cu sau fără fotografie (e.g. fotografie cu benzen). Deși fotografiile ilustrau pur și simplu mărfurile, oamenii au avut tendința de a crede afirmațiile când vedeau o poză. Acest tipar a fost pronunțat în cazul afirmațiilor pozitive referitoare la performanța viitoare a unei mărfi (e.g. marfa va aduce profit), fenomen compatibil cu distorsiunile cognitive de tip „ochelari roz“ care apar în evaluarea evenimentelor din viitor (e.g. Szpunar, Addis și Schacter, 2012). Într-o sarcină de evaluare a unui produs, participanții au avut o probabilitate mai mare de a crede afirmații pozitive despre vinuri (e.g. „Acest vin a fost evaluat ca fiind de calitate“ dacă pe etichete [e.g. Two Quills Pinot Noir (Pinot Noir Două Condeie)] apărea o fotografie neconcludentă reprezentând substantivul nefamiliar din denumirea vinului (e.g. o fotografie cu pene într-o călămară). Acest tipar s-a menținut chiar și când participanților li s-a oferit ocazia de a gusta vinurile cu etichete pe care apărea sau nu o fotografie. Cu alte cuvinte, chiar și cu informații senzoriale suplimentare — gustul vinului — fotografiile aveau un efect de distorsiune asupra judecăților (Carwell, Newman, Garry, Mantonakis și Beckett, 2017).

Într-un experiment care implica amintirea acțiunilor, Carwell, Henkel, Garry, Newman și Foster (2016) au descoperit că oamenii au tendința de a crede că au realizat anumite sarcini (au răspuns că afirmații precum „I-am dat să mănânce acestui animal“ erau adevărate într-un joc care simula o grădină zoologică) dacă vedeau în timpul testului o fotografie neconcludentă. E semnificativ că fotografiile i-au făcut pe oameni să creadă că îndeplinisera o anumită acțiune indiferent dacă au făcut asta sau nu. Acest rezultat este deosebit de surprinzător — deși există unele dovezi că fotografiile pot

distorsiona amintirea acțiunilor recente, fotografiile din acele studii sunt adesea combinate cu variabile ca sugestia, elaborarea și repetiția, toate făcând mai dificilă stabilirea faptului dacă o amintire este sau nu adevărată (Lindsay et al., 2014; pentru o analiză, vezi Garry și Gerrie, 2005). Rezultatele experimentului realizat de Cardwell et al. (2016) demonstrează că până și o expunere scurtă la fotografii neconcludente, în absența altor tehnici de sugestie, poate duce la erori imediate de memorie. Luate laolaltă, aceste rezultate indică faptul că adăugarea de fotografii neconcludente nu doar modelează convingeri referitoare la afirmații generale de cunoaștere, ci fotografiile au un efect sistematic asupra convingerilor în numeroase domenii și pot influența estimările oamenilor în privința prețurilor viitoare ale mărfurilor, calității produselor și judecăților referitoare la propriile acțiuni realizate recent. Cum exercită fotografiile aceste efecte în cazul evaluărilor adevărului?

Mecanisme subiacente senzației de veridicitate

Deși aceste fotografii nu oferă dovezi despre adevărul unei afirmații, există mai multe motive pentru care adăugarea unei fotografii ce are legătură cu afirmația poate influența judecata oamenilor. Expunem aici aceste posibilități și analizăm dovezile empirice disponibile. În primul rând, examinăm posibilitatea ca fotografiile să provoace un efect de distorsiune pentru că sunt intrinsec demne de încredere. Apoi discutăm posibilitatea ca fotografiile să influențeze convingerile printr-un mecanism de fluentă cognitivă, facilitând procesarea semantică a unei afirmații. După care examinăm posibilitatea ca fotografiile să influențeze convingerile pentru că, deși nu sunt concludente, provoacă iluzia dovezilor. Prezentăm dovezi empirice pentru fiecare dintre aceste perspective teoretice.

Fotografiile neconcludente provoacă senzația de veridicitate pentru că sunt intrinsec demne de încredere?

Fotografiile sunt unice sub numeroase aspecte — captează un moment, în general reprezintă evenimente reale și uneori constituie cele mai bune dovezi că un lucru s-a întâmplat cu adevărat (Kelly și Nace, 1994; Strange, Garry, Bernstein și Lindsay, 2011; vezi și Nightingale, Wade și Watson, 2017). Mai mult, oamenii tind să aibă încredere în fotografii și adesea nu pot să-și dea seama dacă o fotografie a fost modificată pentru a induce în eroare (e.g. Nightingale et al., 2017). Astfel, în paradigma senzației de veridicitate, deși fotografiile nu oferă informații concludente pentru o anumită afirmație, ele ar putea totuși să accentueze credința că afirmația este adevărată pentru că fotografiile sunt inerent credibile.

Pentru a examina ipoteza că fotografiile influențează convingerile prin credibilitatea lor intrinsecă, Newman et al. (2012) au examinat dacă efectul senzației de veridicitate este specific fotografiilor sau dacă oamenii sunt influențați și de alte forme de informații care au legătură cu afirmațiile, dar sunt neconcludente. În cadrul unui experiment menit a testa această ipoteză, participanții au văzut nume de celebrități și li s-a cerut să evalueze adevărul afirmației că „Această persoană este vie [moartă]“. În jumătate dintre cazuri, numele celebrităților apăreau alături de informații neconcludente semantice. Pentru un grup de subiecți, informația neconcludentă era o fotografie în care apărea vedeta îndeplinindu-și activitatea, după cum am descris anterior. Pentru celălalt grup de subiecți, informația neconcludentă era o descriere verbală a vedetei îndeplinindu-și activitatea. De exemplu, în loc să vadă o fotografie cu Nick Cave, persoanele din grupul cu informații verbale au văzut o listă de informații semantice care puteau fi extrase cu ușurință din fotografie (e.g. bărbat, brunet, microfon, tamburină; vezi Figura 6.2).

Descoperirea esențială a acestui studiu a fost că, indiferent dacă oamenii au văzut o fotografie cu Nick Cave sau o descriere verbală a lui Nick Cave, aveau o probabilitate mai mare de a crede afirmația „Nick Cave era viu [sau mort]“ în comparație cu cazul în care afirmația apărea fără adăugarea unor informații neconcludente (Newman et al., 2012). Adică, adăugarea de informații semantice, nu neapărat fotografii, a provocat efectul senzației de

veridicitate. Aceste rezultate se încadrează în cercetările mai extinse asupra evaluării adevărului și fluenței cognitive (vezi Schwarz, 2015). De fapt, din ce în ce mai multe studii sugerează că adăugarea de informații asociate conceptual (e.g. fotografii sau cuvinte neconcludente) amplifică procesarea semantică a afirmațiilor, influențând oamenii să creadă că afirmațiile asociate sunt adevărate. Abordăm în continuare această posibilitate.

Fotografiile neconcludente determină senzația de veridicitate printr-o amplificare a fluenței cognitive?

De la un moment la altul, oamenii observă modificări ale propriei procesări cognitive. Indiferent dacă încercăm să ne amintim informații generale sau să ne imaginăm un scenariu, există variații în privința efortului pe care trebuie să-l depunem pentru procesarea informațiilor. Suntem în general conștienți dacă perceperea, înțelegerea sau imaginarea unei informații este simplă și nu necesită efort sau dacă e dificilă și presupune efort (pentru analize, vezi Alter și Oppenheimer, 2009; Jacoby, Kelley și Dywan, 1989; Schwarz, 2015). Când procesarea informației este simplă și se desfășoară ușor, avem tendința să fim de acord (vezi Marsh și Stanley, Capitolul 8; Schwarz și Jalbert, Capitolul 5). Adică, dacă putem extrage cu ușurință din memorie informații relevante, dacă înțelegem repede și generăm rapid imagini mentale în legătură cu o afirmație sau cu o idee avem tendința de a crede că e adevărată. După cum descriu Schwarz și Jalbert în acest volum, simplitatea procesării este interpretată ca indiciu important în evaluarea adevărului, în parte pentru că procesarea simplă este interpretată ca dovadă a numeroase criterii esențiale legate de evaluarea adevărului: informația ușor de procesat este considerată mai coerentă, credibilă și compatibilă cu propriile noastre cunoștințe generale, având probabil un nivel ridicat de consens social și fiind susținută de dovezi semnificative (Schwarz, 2015; Schwarz și Newman, 2017).



bărbat alb
păr lung, drept, negru
muzician
microfon

Figura 6.2. Exemple de informații vizuale și verbale, Newman et al. (2012)

Fotografiile neconcludente ar trebui (teoretic) să faciliteze procesarea conceptuală a unei afirmații

Există mai multe motive pentru care o fotografie neconcludentă (sau cuvinte asociate) pot influența ușurința procesării unei afirmații și îi pot face pe oameni să concluzioneze că o afirmație este adevărată (vezi Alter și Oppenheimer, 2009 pentru o analiză a factorilor care determină fluența cognitivă). În primul rând, o fotografie neconcludentă asociată subiectului ar trebui să ofere un context bogat din punct de vedere semantic pentru evaluarea unei afirmații, facilitând extragerea rapidă a informațiilor legate de subiectul afirmației (vezi Whittlesea, 1993; Wilson și Westerman, 2018). Cu alte cuvinte, în cazul evaluării afirmației „Girafele sunt singurele mamifere ce nu pot să sară”, o fotografie cu o girafă ar trebui să ajute la extragerea mai rapidă a detaliilor (dar probabil neconcludente) despre girafe (vezi Carr, McCauley, Sperber și Parmelee, 1982). Când oamenii pot să extragă rapid informații din memorie (indiferent dacă informațiile sunt relevante sau nu) au tendința de a concluziona că ușurința găsirii acelor informații indică frecvența, familiaritatea și adevărul (vezi conceptele anexe de prioritate semantică și disponibilitate cognitivă; Begg, Anas și Farinacci, 1992; Kelley și Lindsay, 1993; Tversky și Kahneman, 1973; Whittlesea, 1993). În al doilea rând, o fotografie care are legătură cu afirmația ar trebui să-i ajute pe oameni să-și imagineze mai clar despre ce e vorba. Cercetările arată că, dacă oamenii își pot imagina cu ușurință un eveniment ipotetic, au tendința de a concluziona că e mai probabil (Sherman, Cialdini, Schwartzman și Reynolds, 1985; vezi Alter și Oppenheimer, 2009 pentru o analiză). În al treilea rând, stabilirea unei legături între o fotografie și o afirmație accentuează caracterul concret al afirmației, ceea ce amplifică ușurința înțelegerii elementelor acesteia. Dacă informația este scrisă într-un limbaj concret și ușor de înțeles, oamenii au o probabilitate mai mare de a o crede (Hansen și Wanke, 2010; Oppenheimer, 2006). Adică, deși informațiile neconcludente, asociate subiectului, nu constituie dovezi, ele ar trebui să faciliteze procesarea cognitivă în trei moduri esențiale: (1) contribuind la receptarea rapidă a

detaliilor semantice relevante, (2) contribuind la imaginarea mai simplă a afirmației și (3) contribuind la înțelegerea afirmației asociate. Este bine documentat în literatura amplă asupra cogniției faptul că oricare dintre aceste variabile poate determina o mai mare fluență cognitivă și probabilitatea ca oamenii să fie de acord cu o afirmație, ceea ce sugerează că fotografiile neconcludente pot reprezenta o variabilă deosebit de influentă în evaluarea adevărului.

Dovezi experimentale ale influenței pe care fluența cognitivă o are asupra senzației de veridicitate

Care este dovada că fotografiile influențează evaluările adevărului printr-o schimbare a ușurinței de procesare a afirmației? Există mai multe predicții care apar din literatura despre fluența cognitivă. Prima este că o fotografie neconcludentă îi face pe oameni să creadă că o afirmație este adevărată, în măsura în care fotografia facilitează procesarea conceptuală a afirmației. Astfel, fotografiile neconcludente determină senzația de veridicitate când au o legătură semantică cu afirmația, dar nu și în cazul contrar (vezi Alter și Oppenheimer, 2009). Adăugarea de fotografii legate semantic ar trebui să faciliteze procesarea conceptuală a afirmației, iar adăugarea de fotografii fără legătură semantică ar trebui să interfereze cu procesarea afirmației, în comparație cu situația în care nu e prezentă nicio fotografie (Figura 6.3). Pentru a testa această idee, Newman et al. (2015) au folosit paradigma cunoștințelor generale descrise anterior și au mai adus o ajustare metodei: au manipulat relația semantică dintre fotografii și afirmații.

La fel ca în studiile descrise anterior, participanții au văzut afirmații de cultură generală însoțite sau nu de o fotografie și au trebuit să indice dacă acele afirmații erau adevărate sau false. Manipularea esențială din aceste studii a constat în faptul că jumătate dintre subiecți au văzut fotografii care aveau legătură semantică cu afirmațiile (e.g. afirmația „Magneziul este metalul lichid din termometru“ era însoțită de fotografia unui termometru) și jumătate dintre subiecți au văzut fotografii fără legătură semantică cu afirmațiile (e.g. afirmația „Magneziul este metalul lichid din termometru“ era

însoțită de fotografia unei șopârle). Rezultatele acestor studii sunt consecvente interpretării centrate pe rolul fluenței cognitive asupra senzației de veridicitate — în comparație cu situația în care nu apărea nicio fotografie, fotografiile care aveau legătură semantică cu afirmațiile și care ar fi trebuit să faciliteze procesarea conceptuală i-au făcut pe participanți să creadă că afirmația asociată era adevărată, în timp ce fotografiile fără legătură semantică cu afirmația, care ar fi trebuit să interfereze cu procesarea conceptuală a afirmației, i-au făcut pe participanți să creadă că afirmația asociată era falsă, i.e. au provocat apariția unei senzații de falsitate (falsiness effect).

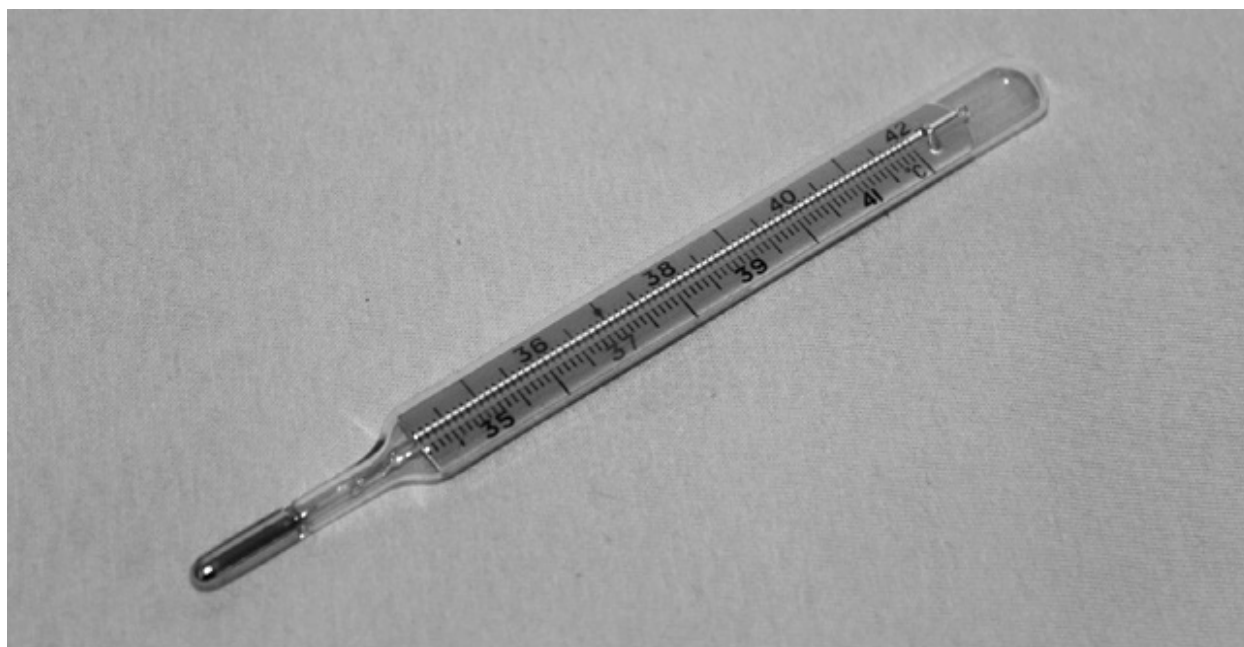


Figura 6.3. Exemple reprezentative de fotografii cu și fără legătură semantică folosite în studiul realizat de Newman et al. (2015)

Sursă: Fotografia unui termometru: Creative Commons License attribution, Da Sal. www.flickr.com/photos/112931986@N07/11646098174. Șopârlă: utilizare gratuită, <https://routespartnership.org/industry-resources/images-video-library/free-use-wildlife-images/reptiles>

O altă predicție a literaturii asupra fluenței este că efectul senzației de veridicitate ar trebui să fie mai accentuat atunci când afirmațiile care apar însoțite de fotografii sunt prezentate printre alte afirmații, care nu sunt însoțite de fotografii. Din ce în ce mai multe cercetări arată că efectele validate ale fluenței sunt mai accentuate în designurile de cercetare intra-subiect decât în cele inter-subiect (e.g. Mere Exposure Effect, Truth Effect, Recognition Memory Effects: Hansen, Dechêne și Wänke, 2008; Westerman, 2008; pentru o analiză, vezi Wänke și Hansen, 2015). Simplu spus, un element important al efectelor bazate pe fluentă este relativa ușurință de procesare a informației. Acest aspect nu aparține exclusiv fluenței. Într-adevăr, schimbările la nivelul experienței au în general un rol în mai mare măsură informativ decât stările stabile (Berelson și Steiner, 1964). Tocmai asta s-a desoperit în paradigmele senzației de veridicitate: efectul senzației de veridicitate este mai accentuat în designurile experimentale care manipulează prezența unei fotografii în cazul studiilor intra-subiect, dar nu și în cele inter-subiect (Newman et al., 2015).

Perspectiva fluenței prezice și faptul că gradul în care fotografiile facilitează procesarea unei afirmații ar trebui să depindă de familiaritatea elementelor esențiale ale afirmației (vezi Parks și Toth, 2006; Whittlesea și Williams, 1998). Adică afirmațiile care conțin elemente ușor de înțeles ar trebui să se situeze mai aproape de un „prag de fluentă” decât afirmațiile care conțin elemente nefamiliare (vezi Carwell, Newman, Garry, Mantonakis și Beckett, 2017; Zhang, Newman și Schwarz, 2018). Așadar, afirmații care includ elemente familiare precum „Vinul Pinot Noir Doi Trandafiri a fost evaluat ca fiind de calitate” ar trebui să beneficieze în mai mică măsură de prezența unei fotografii decât afirmații ca „Vinul Pinot Noir Două Condeie a fost evaluat ca fiind de calitate”. Mai multe descoperiri din literatura asupra senzației de veridicitate sunt compatibile cu aceste predicții. Primele experimente asupra senzației de veridicitate au arătat că senzația de veridicitate a fost mai pronunțată în cazul numelor de vedete nefamiliare, ca Nick Cave, în comparație cu nume de vedete mai familiare, ca Brad Pitt (Newman et al., 2012). Mai recent, Abed, Fenn și Pezdek (2017) au descoperit că fotografiile îi influențează pe oameni să creadă afirmații despre persoane nefamiliare, dar nu și afirmații despre propria persoană, efect determinat poate de diferențele

dintre informațiile relevante care sunt disponibile. Mai mult, Cardwell et al. (2017) au descoperit că măsura în care fotografiile determină o distorsiune a senzației de veridicitate depindea de gradul în care prezența unei fotografii contribuia la înțelegerea elementelor nefamiliare dintr-o afirmație.

Luată laolaltă, aceste efecte sunt consecvente perspectivei fluenței — fotografiile (sau cuvintele) influențează probabil oamenii să creadă prin facilitarea procesării conceptuale a afirmațiilor asociate. Deși aceste studii oferă dovezi convergente în sprijinul perspectivei rolului pe care fluența îl joacă în determinarea senzației de veridicitate, există cel puțin o altă posibilitate: poate că fotografiile influențează judecățile oamenilor în parte pentru că oamenii presupun că experimentatorii, în toate aceste studii diferite, oferă dovezi relevante pe care participanții ar trebui să le folosească pentru a răspunde.

Fotografiile neconcludente produc efectul senzației de veridicitate printr-o iluzie a existenței dovezilor?

Când comunică, oamenii se așteaptă ca emițătorii să ofere doar informații relevante pentru acea conversație și să se abțină să dea mai multe informații decât e necesar (Grice, 1975). Așadar, dacă o fotografie care are legătură cu subiectul, dar nu e concludentă, însoțește o afirmație, oamenii au așteptarea rezonabilă ca fotografia să fie relevantă pentru chestiunea în cauză și ca prezența fotografiei să reprezinte încercarea emițătorului de a oferi dovezi în favoarea afirmației (pentru analize ale efectelor Grice, vezi Schwarz, 1994, 1996). Din această perspectivă, senzația de veridicitate poate fi dată în parte de iluzia că fotografia oferă dovezi și nu de schimbări la nivelul ușurinței procesării determinate de fotografie.

Pentru a separa perspectiva iluziei dovezilor de cea a fluenței, Zhang et al. (2020) au examinat efectele senzației de veridicitate în contextul afirmațiilor comparative. Autorii au modificat materialele din sarcina descrisă anterior de predicție a prețului mărfurilor (Newman, Azad, Lindsay și Garry, 2018), astfel încât (1) fiecare afirmație era prezentată într-un format comparativ și

(2) o jumătate dintre participanți au văzut afirmațiile comparative despre mărfurile ușor de vizualizat (e.g. „Prețul creveților va crește în trei luni mai mult decât prețul trandafirilor“) și cealaltă jumătate a văzut afirmațiile despre mărfuri dificil de vizualizat (e.g. „Prețul betelului va crește în trei luni mai mult decât prețul rasei leghorn“). Într-o treime dintre cazuri, aceste afirmații apăreau alături de o fotografie a subiectului lingvistic (e.g. Crevete pentru grupul cu vizualizare simplă și Betel pentru grupul cu vizualizare dificilă), într-o treime dintre cazuri afirmațiile apăreau alături de o fotografie a referentului lingvistic (e.g. Trandafir pentru grupul cu vizualizare simplă și găini leghorn pentru grupul cu vizualizare dificilă) și într-o treime dintre cazuri afirmațiile nu erau însoțite de nicio fotografie. Utilizarea de afirmații comparative și variația gradului de vizualizare a itemilor le-a permis lui Zhang și colegilor săi să separe perspectiva fluentei de perspectiva iluziei dovezilor din următoarele motive. Dintr-o perspectivă a iluziei dovezilor, și fotografiile subiectului (betel/creveți), și ale referentului (leghorn/trandafiri) sunt relevante din punct de vedere conversațional și conceptual pentru afirmație și pot fi tratate ca dovezi care susțin și afirmațiile cu vizualizare simplă, și afirmațiile cu vizualizare dificilă. Dacă efectul senzației de veridicitate este într-adevăr produs de o iluzie a dovezilor, oricare fotografie ar fi trebuit să influențeze participanții să accepte afirmația ca fiind adevărată și în condiția cu vizualizare simplă, și în cea cu vizualizare dificilă.

Însă, dintr-o perspectivă a fluentei, fotografiile subiectului ar trebui să producă senzația de veridicitate, în timp ce fotografiile referentului ar trebui să determine senzația de fals. Mai mult, efectele de senzație de veridicitate și de senzație de fals ar trebui să depindă de schimbările nivelului de dificultate de procesare produse de fotografii. Cercetarea asupra judecăților comparative arată că o comparație începe cu o evaluare a caracteristicilor subiectului, care apoi sunt puse în relație cu caracteristicile referentului (e.g. Tversky, 1997; Tversky și Gati, 1978). Simplu spus, subiectul este în centrul afirmației și oamenii tind să proceseze subiectul înaintea referentului. Astfel, simplitatea sau dificultatea procesării subiectului ar trebui să aibă consecințe pentru procesare și acceptarea ulterioară a afirmației. Așadar, perspectiva fluentei prezice că, în comparație cu condiția fără fotografii, o fotografie a subiectului unei afirmații comparative ar trebui să determine acceptarea afirmației prin facilitarea procesării subiectului (e.g. prezentarea unei fotografii cu betel pentru afirmația „Prețul betelului va crește în trei luni mai mult decât prețul

rasei leghorn“, mai ales dacă subiectul este altfel dificil de vizualizat fără o fotografie. Prin opoziție, o fotografie a referentului ar trebui să reducă acceptarea afirmației prin faptul că face ca procesarea subiectului (și, în consecință, procesarea afirmației) să fie mai dificilă decât dacă nu ar fi prezentă nicio fotografie (e.g. prezentarea unei fotografii cu trandafiri pentru afirmația „Prețul creveților va crește în trei luni mai mult decât prețul trandafirilor“), mai ales dacă subiectul este ușor de vizualizat fără fotografie.

Într-adevăr, rezultatele obținute de Zhang et al. au fost consecvente acestor predicții bazate pe perspectiva fluenței. Fotografiile subiectului au determinat o acceptare mai mare, iar fotografiile referentului au determinat o reducere a acceptării afirmației comparative, în funcție de schimbările nivelului de dificultate a procesării pe care le-au provocat fotografiile. Cu alte cuvinte, în loc de a fi tratate ca dovezi, fotografiile referentului au acționat mai degrabă ca fotografiile fără legătură semantică folosite în cercetarea anterioară (vezi Newman et al., 2015), determinând un efect de senzație de falsitate. Aceste descoperiri contrazic perspectiva iluziei dovezilor, care prezice că orice fotografie fără legătură cu subiectul va determina apariția senzației de veridicitate.

Luată laolaltă, dovezi convergente din mai multe modalități de manipulare susțin perspectiva fluenței cognitive asupra senzației de veridicitate. Adăugarea de fotografii legate de subiect, dar neconcludente, poate facilita procesarea unei afirmații, influențând oamenii să creadă că o afirmație este adevărată.

Senzația de veridicitate și iluzia familiarității și cunoașterii

Deși în acest capitol ne concentrăm mai ales asupra senzației de adevăr, există din ce în ce mai multe dovezi care susțin că și fotografiile neconcludente influențează evaluările familiarității, ceea ce are implicații

pentru modul în care oamenii își amintesc și estimează propriile cunoștințe generale când citesc o afirmație. Aceste rezultate au o relevanță critică pentru modul în care oamenii evaluează fake news și dezinformarea, așadar descriem aici aceste descoperiri.

Evaluările nivelului de familiaritate sunt adesea influențate de cât de rapid o idee sau o informație asociată este actualizată din memorie (vezi Jacoby et al., 1989). Ceea ce are sens — în comparație cu ceva nou, dacă un lucru a mai fost văzut înainte e mai simplu de amintit și mai ușor de identificat (e.g. Feustel, Shiffrin și Salasoo, 1983; Tulving și Schacter, 1990). Cu toate acestea, senzația de familiaritate poate fi contrafăcută prin variabile din contextul curent care se întâmplă să faciliteze procesarea informației și care totuși n-au nicio legătură cu expunerea anterioară (Whittlesea, 1993). De exemplu, este bine demonstrat faptul că aspecte ale unui test de memorie de recunoaștere care fac itemii testului să pară fluenți îi pot influența pe oameni să susțină că au mai văzut itemii înainte, indiferent dacă i-au mai văzut sau nu (Westerman, 2008; Jacoby et al., 1989). Un alt exemplu este oferit de faptul că, în cazul unui test de recunoaștere, există o probabilitate mai mare ca oamenii să susțină că au examinat un cuvânt-țintă („barcă”) dacă acest cuvânt apăsarea într-o propoziție predictivă semantic în cadrul testului (e.g. „Marea învolburată zdruncina barca.”) decât dacă apăsarea într-o propoziție mai degrabă neutră (e.g. „A economisit și a cumpărat o barcă”; Whittlesea, 1993). Recent, Wilson și Westerman (2018) au demonstrat că adăugarea de fotografii neconcludente la un test de memorie de recunoaștere influențează într-un mod asemănător judecățile mnezice. Fotografiile nu numai că au determinat creșterea vitezei cu care oamenii au identificat itemii testului, ci au și produs iluzia familiarității — făcându-i pe oameni să susțină că au mai văzut cuvintele pe care nu le examinaseră anterior (vezi Brown și Marsh, 2008; fotografii care apar de mai multe ori pot, de asemenea, să ducă la iluzii de familiaritate).

Dat fiind că fotografiile pot provoca iluzii de familiaritate, poate nu e surprinzător că pot și să-i facă pe oameni să supraestimeze senzația de familiaritate pe care o au față de concepte complexe în cadrul unei sarcini de estimare a cunoașterii. Într-o serie de experimente, Cardwell, Lindsay, Förster și Garry (2017) le-au cerut participanților să scoreze cât de pricepuți sunt la diferite procese complexe (e.g. cum se formează curcubeiele). Într-o

jumătate dintre cazuri, participanții au văzut și o fotografie neconcludentă alături de afirmația referitoare la proces (e.g. o fotografie cu fața unui ceas alături de itemul „Cum funcționează ceasurile“). Deși fața de ceas nu oferă nicio informație relevantă în privința mecanismelor ceasului, când oamenii au văzut această fotografie au susținut că știu mai multe despre acest proces. Când Carwell et al. au examinat cunoașterea concretă a acestor procese, persoanele care au văzut fotografiile aveau explicații de o calitate asemănătoare celor ale participanților care nu au văzut o fotografie. În contextul fake news și al dezinformării, astfel de rezultate sunt deosebit de îngrijorătoare și sugerează că fotografiile generice folosite în media pot nu doar să influențeze evaluările cu privire la adevăr, ci s-ar putea și să le provoace oamenilor senzația că știu mai multe sau își amintesc mai multe în legătură cu o anumită afirmație.

Reducerea senzației de veridicitate

Date fiind efectele puternice pe care le au informațiile neconcludente, o întrebare esențială este: cum se poate reduce senzația de veridicitate? Punem această întrebare din trei perspective diferite, luând în calcul (1) condițiile în care oamenii sunt susceptibili la senzația de veridicitate, (2) diferențele individuale în privința susceptibilității la senzația de veridicitate și (3) faptul dacă instrucțiunile sau avertismentele în privința fotografiilor pot influența susceptibilitatea.

În ce condiții sunt oamenii susceptibili la senzația de veridicitate?

Cunoștințe anterioare

Conform literaturii despre fluența cognitivă, oamenii sunt în general mai puțin influențați de experiența de procesare simplă dacă au multe cunoștințe anterioare despre afirmațiile pe care le evaluează (e.g. Dechêne, Stahl, Hansen și Wänke, 2010; Parks și Toth, 2006; Unkelbach, 2007). Acest tipar este valabil și în privința senzației de veridicitate. După cum am descris anterior, efectele senzației de veridicitate sunt în general mai reduse și mai puțin semnificative dacă oamenii evaluează afirmații la care se poate răspunde ușor pe baza cunoștințelor anterioare (Newman et al., 2012). Adică afirmații de cultură generală, cum ar fi „Muntele Everest este cel mai înalt munte“, tind să producă efecte mai mici ale senzației de veridicitate (Newman et al., 2012; despre efecte provocate de familiaritatea substantivelor, vezi și Cardwell et al., 2016, 2017; Zhang et al., 2018). Este demn de remarcat faptul că, deși efectul senzației de veridicitate este redus în aceste condiții, nu este cu totul eliminat (Cardwell et al., 2016, 2017; Newman et al., 2012). Acest efect de atenuare este consecvent cercetărilor asupra efectului de adevăr provocat de repetiții. Descoperirea esențială din literatura asupra efectului de adevăr este că afirmațiile repetate au o probabilitate mai mare de a fi considerate adevărate în comparație cu afirmațiile noi (Bacon, 1979; Begg et al., 1992; Hasher, Goldstein și Toppino, 1977; pentru o analiză, vezi Dechêne et al., 2010). La fel ca senzația de veridicitate, acest efect este mai semnificativ în cazul în care oamenii evaluează afirmații ambigue, dată fiind o mai mare fluență de procesare provocată de expuneri repetate (vezi Dechêne et al., 2010). Însă Fazio et al. (2015) au arătat că tendința de a crede afirmații repetate se menține chiar și dacă, în alte condiții de testare, participanții demonstrează că știu răspunsul corect la afirmații (Fazio et al., 2015; vezi și Henkel și Mattson, 2011). În același domeniu, Unkelbach și Greifeneder (2018) au descoperit efecte semnificative ale fluenței în cazul în care participanților le erau prezentate indicii despre acuratețea afirmațiilor atunci când făceau evaluări ale adevărului. Luate laolaltă, descoperirile din literatura asupra senzației de veridicitate și din literatura asupra efectului de adevăr sugerează că experiența procesării simple este un input semnificativ pentru evaluarea adevărului, în ciuda prezenței cunoștințelor generale și a altor inputuri relevante (vezi Unkelbach și Greifeneder, 2018 pentru o analiză a efectelor combinate ale fluenței și inputurilor declarative).

Valență negativă

Un alt tipar consecvent din literatura asupra senzației de veridicitate este că tendința de a crede afirmații însoțite de fotografii este adesea eliminată complet sau inversată dacă oamenilor li se cere să evalueze afirmații cu valență negativă. De exemplu, în cadrul unei cercetări, când participanților li s-a cerut să evalueze afirmația: „Acest vin a fost considerat de calitate slabă“, prezența unei fotografii fie nu a avut niciun efect, fie a provocat o tendință de a nu crede afirmația (Cardwell et al., 2017). În mod asemănător, în experimentele asupra afirmațiilor referitoare la propriile acțiuni, când oamenilor li s-a cerut să evalueze afirmația: „I-am dat acestui animal hrană nesănătoasă“, o fotografie a animalului i-a făcut pe oameni să nu creadă afirmația (Cardwell et al., 2016). Este semnificativ faptul că în aceste studii a fost găsit efectul tipic al senzației de veridicitate în cazul afirmațiilor pozitive; cu alte cuvinte, exista o probabilitate mai mare ca oamenii să creadă că același vin a fost considerat de calitate înaltă și că i-au dat animalului hrană sănătoasă dacă afirmațiile erau însoțite de o fotografie. De ce efectul tipic al senzației de veridicitate atenuează sau inversează o afirmație cu valență negativă? O literatură amplă arată că senzația de ușurință cognitivă este în sine o experiență pozitivă (Reber, Schwarz și Winkielman, 2004; Winkielman, Schwarz, Fazendeiro și Reber, 2003). Într-adevăr, ușurința procesării amplifică în general evaluările pozitive, dar nu și pe cele negative. De exemplu, experiența fluentei cognitive duce la scoruri mai mari acordate în privința frumuseții și agreabilității, dar nu și a dezagreabilității sau urâteniei (Reber, Winkielman și Schwarz, 1998; Seamon, McKenna și Binder, 1998). Mai mult, experiența procesării ușoare se reflectă în expresii faciale — stimulii procesați cu ușurință duc la activarea mușchiului zigomatic, indicator psihofiziologic al afectului pozitiv (Topolinski, Likowski, Weyers și Strack, 2009; Winkielman și Cacioppo, 2001). Aceste descoperiri pot explica de ce, în contextul unei judecăți evaluative negative, prezența unei fotografii neconcludente are un efect minor asupra evaluărilor adevărului și poate, de fapt, să-i facă pe oameni să considere false afirmațiile cu valență negativă.

Desigur, descoperirile studiului în care participanților li se cerea să evalueze

afirmații prin care o persoană era prezentată drept moartă sau vie par să contrazică această perspectivă. Ne-am putea aștepta ca și judecata referitoare la faptul dacă o vedetă trăiește sau a murit să aibă tot o valență negativă. Dar în studiul despre celebrități vii sau moarte oamenii au făcut judecăți asupra unei serii de nume celebre, astfel încât se poate ca statutul de mort sau viu să fi fost evaluat în funcție de faimă, și nu de valența negativă a morții și pierderii (vezi și Cardwell et al., 2017 pentru o discuție a faptului dacă afirmațiile mort-sau-viu aveau o probabilitate mai mare de a fi interpretate ca fapte mundane, și nu ca afirmații purtătoare de valențe). O mai bună înțelegere a limitărilor efectului valenței este un subiect demn de cercetări viitoare. De exemplu, dacă li s-ar cere oamenilor să se gândească la aspecte condamnabile în comparație cu posibilitatea pierderii unei vedete ar influența măsura în care s-ar găsi un efect al senzației de veridicitate?

O altă întrebare interesantă legată de acest subiect — în special în contextul fake news — este dacă afirmațiile consecvente sau inconsecvente din punct de vedere identitar pot acționa în același mod ca afirmațiile cu valență pozitivă sau negativă. Să luăm în considerare, de exemplu, primul paragraf. Dacă susțin Partidul Democrat și sunt pro-Hillary, adăugarea unei fotografii s-ar putea să-mi influențeze mai puțin credințele decât dacă aș fi un supporter al Partidului Republican. Adică s-ar putea ca senzația de veridicitate să varieze în măsura în care afirmațiile sunt consecvente propriilor convingeri și propriei ideologii politice. Aceasta este o întrebare diferită, căreia i se poate răspunde într-o cercetare viitoare.

Cine este susceptibil la senzația de veridicitate?

Stiluri cognitive

Literatura despre persuasiune și despre schimbarea atitudinii arată că unii oameni au o probabilitate mai mare decât alții de a se baza pe indicatori euristici când evaluează un mesaj. Această diferență individuală este evaluată

în general prin scala Need for Cognition (Necesitate de cogniție) (NFC; Cacioppo și Petty, 1982), care măsoară gradul în care oamenii savurează și se implică în gândire profundă. Cei cu scoruri înalte la NFC au o probabilitate mai mare de a lua în considerare calitatea unui argument sau a unui mesaj, precum și consecvența dovezilor prezentate. Drept urmare, sunt convinși mai degrabă de argumente puternice decât de argumente slabe. Prin opoziție, cei cu scoruri mici la NFC sunt mai puțin interesați de substanța argumentelor și au o probabilitate mai mare de a se baza pe indicatori euristici (cum ar fi statutul emițătorului) și pe informații tangențiale, neconcludente (pentru analize, vezi Petty și Cacioppo, 1986; Priester și Petty, 1995; vezi și Reinhard și Sporer, 2008). Aceste lucruri pot sugera faptul că persoanele cu scoruri mari la NFC ar trebui să aibă o probabilitate mai mică de a fi victime ale senzației de veridicitate, dat fiind că tind să se bazeze mai puțin pe indicatori tangențiali și ar trebui să observe că fotografiile nu sunt concludente. Dar nu descoperim nicio dovadă semnificativă a acestei idei (vezi Newman, Jalbert, Schwarz și Ly, 2020). Ne-am putea aștepta ca alte aspecte implicate în tendința oamenilor de a procesa mai degrabă analitic să reducă efectul senzației de veridicitate. De exemplu, Cognitive Reflection Test (Testul reflecției cognitive) (CRT) măsoară capacitatea persoanei de a depăși reacția intuitivă inițială în favoarea unui răspuns analitic (vezi Pennycook, Fugelsang și Koehler, 2015). Cei cu scoruri mari la CRT ar fi astfel mai puțin influențați de ușurința intuitivă cu care afirmația este procesată și ar avea o probabilitate mai mare de a observa că fotografiile nu sunt concludente.

Vârsta

Este un fapt recunoscut că, pe măsură ce oamenii înaintează în vârstă, capacitatea lor de a-și aminti detalii episodice scade, în timp ce senzația de familiaritate rămâne relativ neafectată (pentru o analiză, vezi Jacoby și Rhodes, 2006). În contextul memoriei, asta înseamnă că adulții mai maturi sunt adesea mai susceptibili la iluzii ale memoriei pentru că tind să se bazeze pe familiaritate, și nu pe amintiri (Jacoby și Rhodes, 2006; despre iluzii ale adevărului, vezi și Skurnik, Yoon, Park și Schwarz, 2005). Însă, în contextul cunoștințelor generale, dovezi recente sugerează că înaintarea în vârstă s-ar

putea să nu afecteze performanța, ci să ofere protecție împotriva distorsiunilor de evaluare a adevărului. Cercetările asupra efectului de adevăr provocat de repetiție arată că, deși adulții maturi și adulții tineri sunt la fel de susceptibili la iluzia de adevăr când judecă afirmații ambigue, adulții maturi prezintă un efect mai mic decât adulții tineri când judecă afirmații mai simple sau mai cunoscute (Brashier, Umanath, Cabeza și Marsh, 2017, vezi și Parks și Toth, 2006). O explicație a acestor descoperiri este că adulții maturi pur și simplu știu mai multe și pot să evalueze mai ușor afirmații bine cunoscute. Însă Brashier și colegii săi au controlat nivelele de cunoaștere ale celor două grupuri prin aplicarea unui test de cunoștințe. De ce sunt adulții maturi mai puțin susceptibili decât adulții tineri la iluzia de adevăr în ce privește afirmațiile de cultură generală? O posibilitate oferită de Brashier et al. este că adulții maturi au rețele de cunoștințe mai bine dezvoltate, ceea ce crește probabilitatea ca ei să aplice cunoștințele relevante în context (vezi Umanath și Marsh, 2014).

Este posibil ca tipare similare să se observe în cazul senzației de veridicitate și al altor efecte determinate de fluentă asupra cunoștințelor generale. Până în prezent a fost realizat doar un studiu inițial care a investigat efectele vârstei asupra susceptibilității la efectul aparenței de adevăr. Într-un studiu recent de la un centru științific local, Derksen, Giroux, Newman și Bernstein (2019) au recrutat participanți cu vârste între 3 și 60 de ani și le-au cerut să ia parte la o versiune ajustată în funcție de vârste a paradigmei senzației de veridicitate (folosind afirmații la care și copiii de trei ani, și adulții maturi ar răspunde cu dificultate, cum ar fi „Ciocănitorea este singura pasăre care poate zbura cu spatele“). Aceste informații nu au indicat diferențe semnificative de mărime a efectului iluziei de adevăr în cazul diferitelor grupe de vârstă (vezi și Thapar și Westerman, 2009). Această cercetare inițială arată că imaginile ar putea juca un rol în ciuda vârstei, deși aici sunt necesare studii ulterioare. Dar această cercetare prezintă și o posibilitate interesantă — așa cum adulții maturi sunt mai puțin susceptibili la efectul expunerii repetate a unor cunoștințe bine știute (Brashier et al., 2017), vârsta poate și proteja împotriva efectului fotografiilor neconcludente în cazul în care afirmațiile pot fi evaluate cu ușurință sau sunt bine cunoscute. Și aceasta este o întrebare empirică interesantă, demnă de cercetări viitoare.

Pot instrucțiunile sau avertismentele să reducă senzația de veridicitate?

Într-o lume în care giganți social media ca Facebook și Google se străduiesc să dezvolte algoritmi de detectare a fotografiilor amăgitoare și a știrilor îndoielnice, există intervenții care se pot face individual pentru ca oamenii să se protejeze și să nu cadă victime efectelor de distorsiune provocate de fotografii? Literatura asupra iluziei de adevăr sugerează că răspunsul este da. Într-adevăr, un avertisment în privința adevărului sau neadevărului afirmațiilor reduce (deși nu elimină) susceptibilitatea oamenilor la iluzia de adevăr determinată de repetiție (Jalbert, Newman și Schwarz, 2020; vezi și Nadarevic și Aßfalg, 2017). Un posibil mecanism al acestei reduceri este faptul că oamenii au examinat afirmațiile cu o privire mai sceptică, depășindu-și senzația de familiaritate. Poate un astfel de mecanism să reducă și efectul fotografiilor neconcludente? Într-un studiu recent, Ly și Newman (2019) au descoperit că, deși un avertisment generic în privința adevărului — „nu vă grăbiți și gândiți-vă critic înainte de a răspunde“ — poate crește acuratețea evaluării afirmațiilor (i.e. capacitatea de a face diferența între afirmațiile adevărate și cele false), nu reduce efectul senzației de veridicitate (tendința de a considera afirmațiile însoțite de fotografii ca fiind adevărate). Acest tipar este compatibil cu rezultatele cercetărilor asupra dezinformării, arătând că avertismentele generice în general nu elimină întru totul efectele dezinformării (e.g. Butler, Zaromb, Lyle și Roediger, 2009; Ecker, Lewndowsky și Tang, 2010).

Senzația de veridicitate nu este eliminată nici de avertismente oferite explicit cu privire la faptul că fotografiile doar ilustrează elemente ale afirmației și nu oferă dovezi în privința acestora, ceea ce sugerează că fotografiile au un efect insidios asupra evaluărilor adevărului (Newman, Garry, Feigenson și Marsh, 2014). Într-adevăr, în cadrul unui studiu, după completarea sarcinii referitoare la predicția prețului mărfurilor, Newman et al. (2018) le-au cerut participanților să spună cum credeau că fotografiile mărfurilor le-au influențat judecățile. Un total de 62% dintre participanți au spus că fotografiile i-au ajutat să înțeleagă sau să-și imagineze subiectul afirmației și 28% au spus că fotografia nu i-a ajutat cu nimic sau i-a încetinit. Doar 10% au spus că au considerat că fotografia oferea credibilitate afirmației. Simplu

spus, oamenii nu au prea înțeles modul în care fotografiile modelează judecățile (vezi Pronin 2007 pentru o trecere în revistă a nereușitei de identificare a distorsiunii). Singura manipulare a instrucțiunilor care a reușit să elimine senzația de veridicitate a fost faptul de a le spune oamenilor să ignore fotografiile (Newman et al., 2014). Această abordare axată pe „ignorare” poate funcționa într-un context experimental, dar are o aplicabilitate limitată într-un cadru natural, în care fotografiile și imaginile tind să capteze atenția și să influențeze timpul petrecut în interacțiunea cu materiale virtuale (e.g. Knobloch et al., 2003).

Senzația de veridicitate și fake news

Formula titlu-fotografie este un șablon pentru știri — adevărate sau false. La prima privire, aceste fotografii decorative par inofensive, dar cercetările trecute aici în revistă arată că semnalizează adevărul și în doar câteva secunde influențează oamenii să creadă. Mărimea efectului senzației de veridicitate variază în funcție de studii și de materiale, dar, printr-un experiment tipic care a folosit afirmații de cultură generală, am estimat recent că mărimea efectului este $d\text{-Cohen} = 0,23$, 95% CI [0,15, 0,30]. Deși participanților din aceste studii li s-a cerut să evalueze adevărul, persoanele care răsfoiesc știri pe internet nu sunt neapărat atente la testarea adevărului sau la avertismentele referitoare la informații false. De fapt, cei care citesc știri online au în medie vizite de doar 2,5 minute pe site-urile ziarelor (Pew Research Center, 2019) și aceste perioade scad și mai mult în cazul în care știrile sunt citite pe dispozitive mobile (Dunaway, Searles, Sui și Paul, 2018). Procesarea relativ relaxată a știrilor, combinată cu campaniile de dezinformare în care știrile și titlurile sunt create pentru a induce în eroare, ar putea însemna că fotografiile produc efecte mai semnificative decât descoperim în general în laborator. Într-adevăr, cercetări recente arată că, dacă oamenii nu sunt avertizați în privința posibilității ca afirmațiile să fie false, efectul iluziei de adevăr aproape se dublează (Jalbert et al., 2020).

Ușurința distribuirii online a informației poate, de asemenea, să ducă la amplificarea influenței fotografiilor. Deși e dificil de estimat „senzația de

veridicitate în mediul natural“ pe baza experimentelor de laborator, dovezi recente sugerează că efectele fotografiilor neconcludente se mențin când participanții sunt testați într-un mediu simulat de social media (Fenn, Ramsay, Kantner, Pezdek și Abed, 2019). În comparație cu afirmațiile care apăreau neînsoțite de fotografii, participanții nu numai că au găsit și informațiile adevărate, și cele false însoțite de fotografii ca fiind mai probabil să fie adevărate, ci au și dat mai multe like-uri informației și au fost mai înclinați să o distribuie pe o platformă simulată de social media. Însă oamenii nu înțeleg prea bine influența pe care fotografiile o au asupra judecății lor (Newman et al., 2018; vezi Pronin, 2007 pentru o trecere în revistă a neresursei de identificare a distorsiunii). De fapt, indivizii tind să creadă că ceilalți sunt mai vulnerabili la fake news și la efectele media decât ei înșiși (Jang și Kim, 2018; Sun, Pan și Shen, 2008). Aceste descoperiri sunt îngrijorătoare, dat fiind că aproximativ 68% dintre americani raportează că primesc știri prin platformele de social media (Matsa și Shearer, 2018) și mult mai mulți sunt expuși întâmplător la știri când utilizează astfel de platforme.

Deși această carte este în esență despre fake news, se poate să nu fie necesar ca un titlu să inducă intenționat în eroare pentru ca fotografiile să promoveze convingeri eronate. De exemplu, fotografiile pot promova convingeri eronate dacă sunt alăturate unor titluri menite a amuza. Într-adevăr, materiale de pe site-uri ca The Onion sau The Borowitz Report, care publică articole satirice despre evenimente fictive, sunt adesea interpretate ca știri reale (Garrett, Bond și Poulsen, 2019). Deși aceste articole sunt scrise într-un limbaj exagerat pentru a sublinia natura satirică a acestora, sunt aproape întotdeauna însoțite de fotografii decorative, și uneori modificate, care au legătură cu evenimentele fictive. Dacă oamenii nu înțeleg caracterul satiric, este probabil ca fotografiile decorative să susțină încrederea în relatări menite a amuza, nu a informa. Mai mult, unele titluri de știri sunt exagerate, deși nu într-un totu false, pentru a atrage atenția cititorilor. Însoțite de o fotografie, astfel de titluri pot părea reale inițial, dar mulți cititori nu dau click și citesc doar titlul, ratând nuanțele din textul materialului (Gabiolkov, Ramachandran, Chaintreau și Legout, 2016). Dat fiind că fotografiile au efecte durabile asupra convingerilor (Fenn et al., 2013) și produc amintiri false (Cardwell et al., 2016; Strange, Garry, Brenstein și Lindsay, 2011), simpla expunere la astfel de combinații titlu-fotografie poate crea și ancora convingeri eronate.

Progresele tehnologice și viitorul fake news

Odată ce progresul tehnologic amplifică accesul digital și răspândirea informației, detectarea fake news pe diferitele platforme social media va fi esențială pentru abordarea potențialului de convingeri eronate. Un procent estimat de 64% dintre americani raportează că articolele fake news le provoacă o confuzie semnificativă în privința evenimentelor curente și aproximativ un sfert dintre americani spun că au distribuit online, intenționat sau neintenționat, articole de dezinformare (Barthel, 2016). Dat fiind că avertismentele tind să fie mai eficiente dacă sunt oferite înainte, și nu după expunere (Jalbert, Newman și Schwarz, 2020), este deosebit de importantă dezvoltarea de algoritmi de identificare înainte de expunere a conținuturilor online de dezinformare și de tip clickbait*****.

Reducerea atenției media care adesea urmează titlurilor false poate, de asemenea, să contribuie la diminuarea convingerilor eronate. Analiza de text în social media arată că materialele de dezinformare tind să reapară de mai multe ori după publicarea inițială, în timp ce faptele reale nu prezintă acest aspect (Shin, Jian, Driscoll și Bar, 2018). Ceea ce este deosebit de îngrijorător, dat fiind că expunerea repetată la informații false provoacă o mai mare familiaritate a articolelor fake news, făcând ca detaliile să fie memorate în timp (Bacon, 1979; Begg et al., 1992; Hasher et al., 1977; pentru o analiză, vezi Dechêne et al., 2010). Acest lucru este adevărat chiar și în cazul în care intenția repetării este corectarea detaliilor false, cum ar fi abordarea „mituri versus fapte” utilizată în general în campanii al căror scop este combaterea dezinformării (e.g. Skurnik et al., 2005; Schwarz, Sanna, Skurnik și Yoon, 2007; vezi și Schwarz și Jalbert, Capitolul 5). Pentru a evita ironia creșterii probabilității ca informațiile false să fie amintite greșit ca fiind adevărate, poate că ar fi mai bine ca dezinformarea să fie ignorată. În același timp, repetarea informației corecte și a faptelor, însoțite de fotografii decorative, ar putea contribui la amintirea faptelor fără promovarea de mituri. Este important să observăm că, deși fotografiile pot să susțină acceptarea și distribuirea dezinformării, ele pot să promoveze și faptele (Fenn et al., 2019). Astfel, pe lângă investigațiile din prezent asupra modurilor în care

fotografiile pot influența judecata, cercetări viitoare asupra aparenței de adevăr ar trebui să exploreze modurile în care fotografiile neconcludente pot fi utilizate pentru corectarea distorsiunilor și pentru combaterea dezinformării.

Rezumat

Încercarea de a face diferența între informațiile acurate și cele false face parte din viața de zi cu zi, indiferent dacă e vorba de o afirmație întâlnită în social media, de o afirmație de la știri sau de o afirmație dintr-un cadru în care miza e mare, cum ar fi la tribunal. Fotografiile apar în fiecare dintre aceste contexte, fie cu scop decorativ, fie ca instrumente de comunicare al căror obiectiv este să capteze atenția sau să ilustreze un mesaj. Însă aceste fotografii nu doar decorează sau ilustrează — chiar și printr-o scurtă expunere, o fotografie aparent inofensivă, care are legătură cu materialul, poate avea efecte semnificative și ample asupra judecății, având nevoie de doar câteva secunde pentru a influența oamenii să creadă și să dea like-uri unor afirmații. Efectul senzației de veridicitate se adaugă unei literaturi extinse asupra fluenței cognitive și constituie încă un exemplu al modului în care disponibilitatea informațiilor relevante din punct de vedere conceptual poate susține procesarea, în același timp influențând judecata. Deși se poate ca oamenii să fie încrezători în capacitatea lor de a face diferența dintre fapte și ficțiune, dintre adevăr și minciună, dintre ce e real și ce e fals, efectul senzației de veridicitate se încadrează în cercetarea mai amplă asupra memoriei martorilor și identificării minciunilor și adevărului, ceea ce arată că evaluările adevărului sunt failibile și vulnerabile la distorsiuni de care adesea nu suntem conștienți. Dar viitorul nu este într-un totu sumbru. Așa cum o fotografie influențează oamenii să creadă afirmații false, poate face și ca faptele să fie memorate. Așadar, cercetările asupra senzației de veridicitate nu ar trebui să se oprească la înțelegerea modului în care judecata poate fi influențată de fotografii, ci ar trebui să afle și cum fotografiile pot fi utilizate ca arme împotriva dezinformării și ca susținătoare ale faptelor în epoca post-

adevăr.

Note

1.Articolul integral: www.nytimes.com/2017/01/18/us/fake-news-hillary-clinton-cameron-harris.html

2.Newman și colegii săi au numit acest efect de distorsionare al fotografiilor neconcludente efect al senzației de veridicitate, împrumutând formula pe care satiristul american Stephen Colbert a folosit-o referindu-se la „adevărul care vine din instinct, nu din cărți“. Vezi www.colbertnation.com/the-colbert-report-videos/24039/october-17-2005/the-word-truthiness

Referințe

Abed, E., Fenn, E. și Pezdek, K. (2017). Photographs elevate truth judgments about less well-known people (but not yourself). *Journal of Applied Research in Memory and Cognition*, 6, 203–209.

Alter, A.L. și Oppenheimer, D.M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, 13, 219–235. doi:10.1177/1088868309341564

Bacon, F.T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 241–252. doi: 10.1037/0278-7393.5.3.241

Barthel, M. (15 decembrie 2016). Many Americans believe fake news is

sowing confusion. Pew Research Center, Journalism and Media. Accesat la www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/#fn-59275-1

Begg, I.M., Anas, A. și Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, 121 (4), 446.

Berelson, B. și Steiner, G.A. (1964). *Human behavior: An inventory of scientific findings*. Oxford, Marea Britanie: Harcourt, Brace & World.

Brashier, N.M., Umanath, S., Cabeza, R. și Marsh, E.J. (2017). Competing cues: Older adults rely on knowledge in the face of fluency. *Psychology and Aging*, 32 (4), 331.

Brown, A.S. și Marsh, E.J. (2008). Evoking false beliefs about autobiographical experience. *Psychonomic Bulletin & Review*, 15 (1), 186–190.

Butler, A.C., Zaromb, F.M., Lyle, K.B. și Roediger III, H.L. (2009). Using popular films to enhance classroom learning: The good, the bad, and the interesting. *Psychological Science*, 20 (9), 1161–1168.

Cacioppo, J.T. și Petty, R.E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42 (1), 116.

Cardwell, B.A., Henkel, L.A., Garry, M., Newman, E.J. și Foster, J.L. (2016). Nonprobative photos rapidly lead people to believe claims about their own (and other people's) pasts. *Memory & Cognition*, 44, 883–896.

Cardwell, B.A., Lindsay, D.S., Förster, K. și Garry, M. (2017). Uninformative photos can increase people's perceived knowledge of complicated processes. *Journal of Applied Research in Memory and Cognition*, 6, 244–252.

Cardwell, B.A., Newman, E.J., Garry, M., Mantonakis, A. și Beckett, R. (2017). Photos that increase feelings of learning promote positive evaluations. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition, 43 (6), 944.

Carney, R.N. și Levin, J.R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14, 5–26.

Carr, T.H., McCauley, C., Sperber, R.D. și Parmelee, C.M. (1982). Words, pictures, and priming: On semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception and Performance*, 8 (6), 757–777.

Dechêne, A., Stahl, C., Hansen, J. și Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, 14, 238–257.

Derksen, D.G., Giroux, M.E., Newman, E.J. și Bernstein, D. (iunie 2019). Truthiness bias from 3–60 years. Poster prezentat la conferința bianuală a Society for Applied Research on Memory and Cognition. Cape Cod, MA.

Dunaway, J., Searles, K., Sui, M. și Paul, N. (2018). News attention in a mobile era. *Journal of Computer-Mediated Communication*, 23 (2), 107–124.

Ecker, U.K., Lewandowsky, S. și Tang, D.T. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition*, 38 (8), 1087–1100.

Fazio, L.K., Brashier, N.M., Payne, B.K. și Marsh, E.J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General*, 144, 993–1002.

Fenn, E., Newman, E.J., Pezdek, K. și Garry, M. (2013). The effect of non-probative photographs on truthiness persists over time. *Acta Psychologica*, 144 (1), 207–211.

Fenn, E., Ramsay, N., Kantner, J., Pezdek, K. și Abed, E. (2019). Nonprobative photos increase truth, like, and share judgments in a simulated social media environment. *Journal of Applied Research in Memory and Cognition*, 8 (2), 131–138.

Feustel, T., Shiffrin, R. și Salasoo, A. (1983). Episodic and lexical contributions to the repetition effect in word recognition. *Journal of Experimental Psychology: General*, 112, 309–346.

Gabielkov, M., Ramachandran, A., Chaintreau, A. și Legout, A. (iunie 2016). Social clicks: What and who gets read on Twitter? În *Proceedings of the 2016 ACM SIGMETRICS international conference on measurement and modeling of computer science* (pp. 179–192).

Garrett, R.K., Bond, R. și Poulsen, S. (16 august 2019). Too many people think satirical news is real. *The Conversation*. Accesat la <https://theconversation.com/too-many-people-think-satirical-news-is-real-121666>

Garry, M. și Gerrie, M.P. (2005). When photographs create false memories. *Current Directions in Psychological Science*, 14 (6), 321–325.

Grice, H.P. (1975). Logic and conversation. În P. Cole și J.L. Morgan (coord.), *Syntax and semantics*, vol. 3: Speech acts (pp. 41–58). New York: Academic Press.

Hansen, J., Dechêne, A. și Wänke, M. (2008). Discrepant fluency increases subjective truth. *Journal of Experimental Social Psychology*, 44 (3), 687–691.

Hansen, J. și Wanke, M. (2010). Truth from language and truth from fit: The impact of linguistic concreteness and level of construal on subjective truth. *Personality and Social Psychology Bulletin*, 36, 1576–1588. doi: 10.1177/0146167210386238

Hasher, L., Goldstein, D. și Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16, 107–112.

Henkel, L.A. și Mattson, M.E. (2011). Reading is believing: The truth effect and source credibility. *Consciousness and Cognition*, 20 (4), 1705–1721.

Jacoby, L.L., Kelley, C.M. și Dywan, J. (1989). Memory attributions. În H.L. Roediger III și F.I.M. Craik (coord.), *Varieties of memory and consciousness*:

Essays in honour of Endel Tulving (pp. 391–422). Hillsdale, NJ: Erlbaum.

Jacoby, L.L. și Rhodes, M.G. (2006). False remembering in the aged. *Current Directions in Psychological Science*, 15 (2), 49–53.

Jalbert, M., Newman, E.J. și Schwarz, N. (2019). Only half of what I tell you is true: Expecting to encounter falsehoods reduces illusory truth. Manuscris în curs de revizuire.

Jang, S.M. și Kim, J.K. (2018). Third person effects of fake news: Fake news regulation and media literacy interventions. *Computers in Human Behavior*, 80, 295–302.

Kelley, C.M. și Lindsay, D.S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, 32, 1–24. doi: 10.1006/jmla.1993.1001

Kelly, J.D. și Nace, D. (1994). Knowing about digital imaging and believing news photographs. *Visual Communication Quarterly*, 18 (1), 4–5.

Knobloch, S., Hastall, M., Zillmann, D. și Callison, C. (2003). Imagery effects on the selective reading of Internet newsmagazines. *Communication Research*, 30 (1), 3–29.

Lindsay, D.S., Hagen, L., Read, J.D., Wade, K.A. și Garry, M. (2004). True photographs and false memories. *Psychological Science*, 15, 149–154.

Ly, D.P., și Newman, E.J. (aprilie 2019). Thinking more does not protect people from truthiness. Poster prezentat la Experimental Psychology Conference, Victoria University of Wellington, New Zealand.

Matsa, K.E. și Shearer, E. (2018). News use across social media platforms 2018. Pew Research Center, Journalism and Media. Accesat la www.journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/

Marsh, E. și Stanley, M. (2020). False beliefs: byproducts of an adaptive

knowledge base? În R. Greifeneder, M. Jaffé, E.J. Newman și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 202–225). Londra, Marea Britanie: Routledge.

Nadarevic, L. și Aßfalg, A. (2017). Unveiling the truth: Warnings reduce the repetition-based truth effect. *Psychological Research*, 81, 814–826.

Newman, E.J., Azad, T., Lindsay, D.S. și Garry, M. (2018). Evidence that photos promote rosiness for claims about the future. *Memory & Cognition*, 46 (8), 1223–1233.

Newman, E.J., Garry, M., Bernstein, D.M., Kantner, J. și Lindsay, D.S. (2012). Nonprobative photos (or words) promote truthiness. *Psychonomic Bulletin & Review*, 19, 969–974. doi: 10.3758/s13423-012-0292-0

Newman, E.J., Garry, M., Feigenson, N. și Marsh, E.J. (mai 2014). To what extent can jury instructions protect people from truthiness? Prezentare susținută la 26th Association for Psychological Science Annual Convention, San Francisco, CA.

Newman, E.J., Garry, M., Unkelbach, C., Bernstein, D.M., Lindsay, D.S. și Nash, R.A. (2015). Truthiness and falsiness of trivia claims depend on judgmental contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41, 1337–1348. doi: 10.1037/xlm0000099

Newman, E.J., Jalbert, M.C., Schwarz, N. și Ly, D.P. (2020). Truthiness, the illusory truth effect, and the role of need for cognition. *Consciousness and Cognition*, 78, 102866.

Nightingale, S.J., Wade, K.A. și Watson, D.G. (2017). Can people identify original and manipulated photos of real-world scenes? *Cognitive Research: Principles and Implications*, 2 (1), 30.

Oppenheimer, D.M. (2006). Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 20 (2), 139–156.

Parks, C.M. și Toth, J.P. (2006). Fluency, familiarity, aging, and the illusion of truth. *Aging, Neuropsychology, and Cognition*, 13 (2), 225–253.

Pennycook, G., Fugelsang, J.A. și Koehler, D.J. (2015). Everyday consequences of analytic thinking. *Current Directions in Psychological Science*, 24 (6), 425–432.

Petty, R.E. și Cacioppo, J.T. (1986). The elaboration likelihood model of persuasion. În *Communication and persuasion* (pp. 1–24). New York, NY: Springer.

Pew Research Center. (9 iulie 2019). Newspapers fact sheet. Pew Research Center, Journalism and Media. Accesat la www.journalism.org/fact-sheet/newspapers/

Politico Staff. (1 noiembrie 2017). The social media ads Russia wanted Americans to see. Politico. Accesat la www.politico.com/story/2017/11/01/social-media-adsrussia-wanted--americans-to-see-244423

Priester, J.R. și Petty, R.E. (1995). Source attributions and persuasion: Perceived honesty as a determinant of message scrutiny. *Personality and Social Psychology Bulletin*, 21 (6), 637–654.

Pronin, E. (2007). Perception and misperception of bias in human judgment. *Trends in Cognitive Sciences*, 11, 37–43.

Reber, R., Schwarz, N. și Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8 (4), 364–382.

Reber, R., Winkielman, P. și Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, 9, 45–48.

Reinhard, M.A. și Sporer, S.L. (2008). Verbal and nonverbal behaviour as a basis for credibility attribution: The impact of task involvement and cognitive capacity. *Journal of Experimental Social Psychology*, 44 (3), 477–488.

Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation. În M. Zanna (coord.), *Advances in experimental social psychology* (Vol. 26, pp. 123–162). San Diego, CA: Academic Press.

Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods, and the logic of conversation*. Hillsdale, NJ: Erlbaum.

Schwarz, N. (2015). Metacognition. În E. Borgida și J.A. Bargh (coord.), *APA handbook of personality and social psychology: Attitudes and social cognition*. Washington, DC: APA.

Schwarz, N. și Jalbert, M. (2020). When (fake) news feels true: Intuitions of truth and the acceptance and correction of misinformation. În R. Greifeneder, M. Jaffé, E.J. Newman și N. Schwarz (coord.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 113–139). Londra, Marea Britanie: Routledge

Schwarz, N. și Newman, E.J. (2017). Psychological science agenda, august 2017. *Psychological Science*.

Schwarz, N., Sanna, L.J., Skurnik, I. și Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology*, 39, 127–161.

Seamon, J.G., McKenna, P.A. și Binder, N. (1998). The mere exposure effect is differentially sensitive to different judgment tasks. *Consciousness and Cognition*, 7 (1), 85–102.

Sherman, S.J., Cialdini, R.B., Schwartzman, D.F. și Reynolds, K.D. (1985). Imagining can heighten or lower the perceived likelihood of contracting a disease: The mediating effect of ease of imagery. *Personality and Social Psychology Bulletin*, 11, 118–127. doi: 10.1177/0146167285111011

Shin, J., Jian, L., Driscoll, K. și Bar, F. (2018). The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior*, 83, 278–287.

Skurnik, I., Yoon, C., Park, D.C. și Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, 31 (4), 713–724.

Strange, D., Garry, M., Bernstein, D.M. și Lindsay, D.S. (2011). Photographs cause false memories for the news. *Acta Psychologica*, 136 (1), 90–94.

Sun, Y., Pan, Z. și Shen, L. (2008). Understanding the third-person perception: Evidence from a meta-analysis. *Journal of Communication*, 58 (2), 280–300.

Szpunar, K.K., Addis, D.R., și Schacter, D.L. (2012). Memory for emotional simulations: Remembering a rosy future. *Psychological Science*, 23 (1), 24–29.

Thapar, A. și Westerman, D.L. (2009). Aging and fluency based illusions of recognition memory. *Psychology and Aging*, 24, 595–603.

Topolinski, S., Likowski, K.U., Weyers, P. și Strack, F. (2009). The face of fluency: Semantic coherence automatically elicits a specific pattern of facial muscle reactions. *Cognition and Emotion*, 23 (2), 260–271.

Tulving, E. și Schacter, D.L. (1990). Priming and human memory systems. *Science*, 247, 301–306.

Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327–352.

Tversky, A. și Gati, I. (1978). Studies of similarity. În E. Rosch și B. Lloyd (coord.), *Cognition and categorization* (pp. 79–98). Hillsdale, NJ: Erlbaum.

Tversky, A. și Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5 (2), 207–232.

Umanath, S. și Marsh, E.J. (2014). Understanding how prior knowledge influences memory in older adults. *Perspectives on Psychological Science*, 9 (4), 408–426.

Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation

of processing fluency in judgments of truth. *Psychological Science*, 20, 135–138. doi: 10.1037/0278-7393.33.1.219

Unkelbach, C. și Greifeneder, R. (2018). Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology*, 79, 78–86.

Waddill, P.J. și McDaniel, M.A. (1992). Pictorial enhancement of text memory: Limitations imposed by picture type and comprehension skill. *Memory & Cognition*, 20 (5), 472–482.

Wänke, M. și Hansen, J. (2015). Relative processing fluency. *Current Directions in Psychological Science*, 24 (3), 195–199.

Westerman, D.L. (2008). Relative fluency and illusions of recognition memory. *Psychonomic Bulletin & Review*, 15 (6), 1196–1200.

Whittlesea, B.W.A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1235–1253. doi: 10.1037/0278-7393.19.6.1235

Whittlesea, B.W. și Williams, L.D. (1998). Why do strangers feel familiar, but friends don't? A discrepancy-attribution account of feelings of familiarity. *Acta psychologica*, 98 (2–3), 141–165.

Wilson, J.C. și Westerman, D.L. (2018). Picture (im)perfect: Illusions of recognition memory produced by photographs at test. *Memory & Cognition*, 46 (7), 1210–1221.

Winkielman, P. și Cacioppo, J.T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, 81 (6), 989.

Winkielman, P., Schwarz, N., Fazendeiro, T. și Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. *The Psychology of Evaluation: Affective Processes in Cognition and Emotion*, 189, 217.

Zhang, L., Newman, E.J. și Schwarz, N. (2020). When photos backfire: Truthiness and falsiness effects in comparative judgements. Manuscris în curs de revizuire.

Capitolul 7. O fi adevărat sau sunt doar fake news?

Noi perspective asupra tendinței negativiste în judecarea adevărului

Mariela E. Jaffé și Rainer Greifeneder

Introducere

Care dintre următoarele două afirmații considerați că e adevărată: (1) 80% dintre căsătorii durează zece ani sau mai mult sau (2) 20% dintre persoanele care se căsătoresc divorțează în primii zece ani? Care e mai probabil să fie fake news? S-ar putea ca un cititor să ajungă la concluzia că autorii au făcut o greșeală — ambele afirmații sunt identice în ce privește conținutul, așa că nu se poate decât ca ambele să fie adevărate sau false. Într-adevăr, din punct de vedere logic, afirmațiile sunt identice. Cu toate astea, când indivizilor li se cere să evalueze adevărul sau falsitatea acestora, estimările variază în funcție de faptul dacă citesc același lucru formulat într-o versiune pozitivă (1) sau negativă (2). În speță, cercetarea a arătat că afirmația încadrată negativ (2), centrată pe rata divorțului, are o probabilitate mai mare de a fi considerată adevărată în comparație cu versiunea încadrată pozitiv (1), centrată pe durata căsătoriei (Hilbig, 2009). Această distorsionare a fost numită distorsiunea

negativității (negativity bias) în evaluarea adevărului (Hilbig, 2009; 2012a, 2012b) și descrie tendința de a considera adevărată o afirmație încadrată negativ în comparație cu o afirmație identică cu prima în ce privește conținutul, dar încadrată pozitiv. În acest capitol explorăm în detaliu mecanismele psihologice ale acestei distorsiuni și indicăm faptul că e posibil ca aici să fie vorba de mai multe lucruri decât se credea până acum. Trecem în revistă cercetările realizate asupra acestei distorsiuni și analizăm apoi cum și când distorsiunea negativității influențează (sau nu) și în cele din urmă face ca valoarea de adevăr atribuită de individ să încline într-o direcție sau alta. Acest rezumat pune în lumină maleabilitatea evaluărilor adevărului și oferă informații despre modul în care formatul comunicării poate face informația să pară foarte convingătoare, deși conținutul său este fals sau contrafăcut.

Adevăr sau post-adevăr? O discuție în continuă desfășurare

Faptul de a spune adevărul este esențial în interacțiunile noastre cotidiene. Dacă adresăm cuiva o întrebare, în general ne așteptăm ca persoana să ofere informația pe care o consideră adevărată. Această convingere este o necesitate pragmatică pentru funcționarea comunicării umane: când se angajează în interacțiuni interpersonale, este de așteptat ca indivizii implicați să respecte așa-numitul principiu al cooperării (Grice, 1975). Acest principiu al cooperării presupune faptul ca indivizii „să contribuie la conversație în modul adecvat stadiului în care are loc, în scopul sau direcția discuției în care sunt implicați” (Grice, 1975, p. 45). Acest principiu este rezumatul mai multor sub-principii, dintre care unul este maximul calității, conform căruia persoana nu ar trebui să spună ceva ce crede că e fals și nu ar trebui să spună ceva pentru care nu are dovezi adecvate. Pe scurt, ar trebui să spunem adevărul, iar celălalt se poate aștepta să i se spună adevărul.

Deși adevărul pare esențial pentru interacțiunile umane, post-adevărul (post-truth) a fost ales drept cuvânt al anului 2016, definit ca „referindu-se sau

denotând circumstanțe în care faptele obiective sunt mai puțin influente în privința modelării opiniei publice decât apelurile la emoții și credințe personale (Oxford Dictionary, 2016). Autorii dicționarului au susținut că frecvența utilizării acestui cuvânt a explodat în contextul referendumului din 2016 din Marea Britanie și al alegerilor prezidențiale din SUA și a fost astfel asociat cu sintagma specifică de „politică post-adevăr”. Reflectă presupunerea generală că trăim într-o epocă „în care conceptul [de adevăr] a devenit neimportant sau irelevant”, în special în contextul politic (Oxford Dictionary, 2016). Matthew Norman, jurnalist la The Independent, este de acord și observă că „adevărul s-a depreciat atât de mult, încât ceea ce a fost odată etalon al dezbaterii politice este acum total lipsit de valoare” (Norman, 2016) — și, în locul adevărului, așa-zisele fapte alternative sunt prezentate de candidați și politicieni.

Într-o lume a politicii post-adevăr, a faptelor alternative și a incertitudinii, putem spune că adevărul și falsul sunt mai dificil de diferențiat decât oricând până acum. The Guardian a sugerat că este esențială combaterea fake news prin „verificarea faptelor” și că jurnalismul trebuie să se priceapă mai bine la răspândirea faptelor pentru a le arăta oamenilor ce se întâmplă de fapt (Jackson, 2017). Cu toate acestea, pentru a putea convinge indivizii de adevărul acestor fapte și pentru a le oferi avertismente credibile împotriva fake news, trebuie mai întâi să înțelegem modul în care indivizii decid dacă să creadă că o informație este adevărată sau falsă. În acest capitol ne concentrăm asupra unui anumit aspect al acestei chestiuni: distorsiunea negativității în evaluarea adevărului.

Cum evaluează indivizii adevărul?

Grice (1975) a presupus că indivizii pornesc de la premisa că ceilalți spun adevărul (vezi și Marsh și Stanley, volumul de față). Conform acestui principiu fundamental, oamenii supraestimează sinceritatea celorlalți, fenomen cunoscut ca distorsiunea de adevăr (truth bias) (Burgoon, Blair și

Strom, 2008; Reinhard, Greifeneder și Scharmach, 2013) și care a fost amplu consemnat în numeroase contexte (e.g. Anolli, Balconi și Ciceri, 2003; Buller, Burgoon, White și Ebescu, 1994; Vrij și Mann, 2001).

Dacă însă indivizii nu sunt dispuși să pornească de la premisa sincerității, pot apela la strategii pentru a decide dacă o informație este adevărată sau falsă. Dunwoody (2009), de exemplu, a sintetizat că principiile de coerență și corespondență (C&C — coherence & correspondence, Hammond, 1996, 2007) pot fi criterii de judecare a valorii de adevăr a unei afirmații, convingeri sau judecăți. Teoria adevărului potrivit criteriului corespondenței susține că o convingere poate fi considerată adevărată în măsura în care corespunde faptelor (Dunwoody, 2009, p. 117). Ca exemplu de testare a corespondenței, adevărul prognozei meteo poate fi verificat prin comparație cu realitatea (Hammond, 1996, p. 95). De exemplu, dacă se anunța ploaie și chiar ploua, predicția a fost adevărată; dacă nu ploua, predicția a fost falsă. Însă această abordare mai degrabă filosofică a adevărului a fost criticată de pragmatişti și idealişti în secolul XIX (Dunwoody, 2009) și au fost ridicate obiecții în privința noțiunii de fapt. Această critică a favorizat o abordare alternativă, teoria adevărului potrivit criteriului coerenței, conform căreia adevărul este judecat în funcție de consecvența convingerilor (Hammond, 1996). Pentru a testa coerența sau raționalitatea, se poate verifica dacă argumentele care duc la o anumită concluzie îndeplinesc criteriile consecvenței logice sau matematice (Hammond, 1996, p. 95). Deși coerența nu garantează singură adevărul (Dunwoody, 2009), ambele teorii ale adevărului sunt relevante pentru modul în care indivizii pot ajunge la o judecată în ce privește adevărul unei afirmații — verifică dacă afirmația corespunde faptelor pe care le știu și/sau dacă este coerentă credințelor lor.

Această abordare bazată pe teoria coerenței și pe teoria corespondenței poate fi transpusă în procese distincte pe care indivizii le pot aplica în evaluarea adevărului. Reder (1982), de exemplu, susține că o afirmație poate fi verificată fie găsim direct în memorie ceva asemănător conținutului întrebării (abordare care poate fi considerată paralelă principiului corespondenței), fie evaluând veridicitatea afirmației (abordare care poate fi considerată paralelă principiului coerenței). Reder (1982) susține, în continuare, că în viața cotidiană este improbabil să fie stocate în memorie toate faptele sau majoritatea faptelor despre care oamenii sunt întrebați, ceea ce face ca

evaluările bazate pe veridicitate să fie utilizate cel mai frecvent.

Dar cum evaluează și judecă oamenii veridicitatea unui mesaj sau a unei informații? Și mai ales când afirmațiile sunt ambigue și indivizii nu au multe cunoștințe sau convingeri puternice (Dechêne, Stahl, Hansen și Wanke, 2010)? Cel puțin trei surse de informații pot constitui baza pe care poate fi evaluată veridicitatea: indicii despre sursa afirmației (e.g. nivelul de competență al sursei în ce privește subiectul), attributele contextului în care e prezentată (e.g. la o conferință științifică) și attribute ale afirmației propriu-zise (vezi Dechêne et al., 2010).

În special attributele afirmației propriu-zise pot influența direct veridicitatea afirmației și, făcând-o să pară logică, determină și amplificarea coerenței informației. Attributele afirmației propriu-zise pot include repetiția (vezi, de exemplu, Stanley, Yang și Marsh, 2019), dar și recunoașterea, familiaritatea, experiența de procesare sau fluența (e.g. Reber și Schwarz, 1999; pentru o trecere în revistă, vezi Reber și Greifeneder, 2017) și referințele mentale (Unkelbach și Koch, 2019; Unkelbach și Rom, 2017).

În acest capitol ne concentrăm asupra unui atribut al afirmației propriu-zise, și anume cadrul conceptual al acesteia, detaliat în cele ce urmează. Conform lui Hilbig (2009), definim aici cadrul conceptual ca denotând ceva pozitiv sau negativ. În cercetările asupra distorsiunii negativității, cadrul pozitiv sau negativ al unei afirmații este manipulat pentru a descrie ceva pozitiv sau ceva negativ (Hilbig, 2009). De exemplu, când vorbim despre situația căsătoriilor, ne putem concentra asupra (1) procentului de căsătorii care durează zece ani sau mai mult sau (2) asupra procentului de căsătorii care sfârșesc prin divorț în primii zece ani. Cercetările asupra distorsiunii negativității în privința evaluărilor adevărului au observat în general că indivizii folosesc cadrarea afirmației propriu-zise ca indiciu al veridicității (Hilbig, 2009, 2012b, 2012a). Ceteris paribus, indivizii percep informația cadrată negativ (2) ca având o probabilitate mai mare de a fi adevărată decât informația cadrată pozitiv (1).

Distorsiunea negativității în evaluarea

adevărului

Distorsiunea negativității în evaluarea adevărului (Hilbig, 2009) este consecvență unui principiu care poate fi observat în general și conform căruia exemplele negative tind să aibă o influență mai mare decât cele pozitive (Baumeister, Bratslavsky, Finkenauer și Vohs, 2001). Acest principiu răul-este-mai-puternic-decât-binele se aplică în cazul întâmplărilor cotidiene, al evenimentelor majore de viață (e.g. traume), al relațiilor personale apropiate, al tiparelor de interacțiune socială, al interacțiunilor interpersonale și al proceselor de învățare. Se presupune că acest principiu general provine din efectul potențial defavorabil al informațiilor și evenimentelor negative (în comparație cu cele pozitive) asupra supraviețuirii organismului. Mai mult, se pare că oamenii sunt mai motivați să evite definițiile negative de sine decât să încerce să obțină unele pozitive, astfel că informațiile negative sunt considerate a avea o relevanță mai mare și sunt procesate în mai mare profunzime decât cele pozitive (Baumeister et al., 2001).

Distorsiunea generală a negativității se referă specific și la evaluările adevărului: mesajele care sunt echivalente din punct de vedere formal sunt în mai mare măsură considerate adevărate dacă sunt cadrate negativ decât dacă sunt cadrate pozitiv (Hilbig, 2009). Pot fi invocate mai multe perspective pentru a explica această distorsiune. De exemplu, exemplele negative atrag mai mult atenția (Pratto și John, 1991) și sunt percepute ca fiind în mai mare măsură informative (Peeters și Czapinski, 1990), întrucât sunt mai rare și mai amenințătoare (Dijksterhuis și Aarts, 2003; Lewicka, Czapinski și Peeters, 1992; Peeters și Czapinski, 1990). Dat fiind că sunt considerate a avea o relevanță mai mare, informațiile negative (rele) sunt procesate mai în profunzime decât cele pozitive (bune) (Baumeister et al., 2001). Procesarea mai detaliată, mai atentă sau repetată poate astfel să amplifice puterea de convingere a mesajului (Petty și Brinol, 2008; Shiv, Britton și Payne, 2004).

În ceea ce privește evaluările adevărului, Hilbig (2012a) susține că un cadru negativ ar putea nu numai să ducă la o procesare mai profundă, ci și la creșterea nivelului de atenție sau activare, care, la rândul său, permite o mai simplă actualizare a cunoștințelor relevante sau generare a dovezilor.

Specific, cadrul negativ determină o reactualizare mai fluidă sau o generare mai rapidă a informațiilor relevante, ceea ce duce la o procesare mai fluentă și este asociată cu o mai mare veridicitate evaluată subiectiv (Hilbig, 2012a, p. 39). După cum arată Schwarz și Jalbert (volumul de față), experiența fluidă de procesare poate fi utilizată ca bază alternativă pentru evaluarea veridicității informației. În speță, când evaluează credibilitatea sau coerența, indivizii se pot baza pe fluentă pentru a răspunde la întrebări precum: „Este asta compatibilă cu alte credințe ale mele?” sau „Are coerență internă?” (vezi Schwarz și Jalbert, volumul de față).

Rafinarea perspectivei asupra distorsiunii negativității

Un studiu realizat anterior (Hilbig, 2012b) a descris distorsiunea negativității ca pe o distorsiune de răspuns generală. Utilizând un model de procesare polinomial de tip arbore, Hilbig (2012b) a găsit dovezi ale faptului că o distorsiune în ce privește răspunsurile, și nu cunoștințele diferite, explică efectul de cadrare. Dacă au cunoștințe insuficiente, indivizii au o probabilitate mai mare de a presupune că o afirmație statistică încadrată negativ este „adevărată”, decât o afirmație (cu sens identic) încadrată pozitiv.

Acest capitol oferă o perspectivă amplă asupra distorsiunii negativității în cazul evaluărilor adevărului. Am investigat prin propriile cercetări circumstanțele cu potențial de moderare a apariției unei distorsiuni a negativității în evaluarea adevărului. Într-un prim studiu, am utilizat opt afirmații corecte despre femeile din statele care vorbesc germană. Afirmațiile erau despre sănătate, bunăstare și siguranță socială. Participanții au citit fie o versiune negativă, fie una pozitivă a afirmațiilor și apoi au evaluat veridicitatea fiecăreia. Spre surpriza noastră, și contrar dovezilor raportate de Hilbig (2009), o primă analiză a adus în prim-plan distorsiunea pozitivității; adică afirmațiile cadrate pozitiv aveau o probabilitate mai mare de a fi considerate adevărate în comparație cu cele cadrate negativ (Jaffé și

Greifeneder, 2019). Pe fundalul acestei dovezi, se pare că, în privința distorsiunii negativității, sunt mai multe lucruri decât se credea până acum care încă trebuie descoperite. Acest capitol rezumă descoperirile în privința a trei posibile variabile moderatoare: așteptările indivizilor, sursa negativității și distanța psihologică față de conținutul afirmațiilor. Mai mult, acest capitol aduce în discuție și alte variabile care pot avea o influență asupra distorsiunilor pe care indivizii le manifestă când evaluează adevărul.

Potențiale variabile moderatoare ale distorsiunii negativității în evaluarea adevărului

Așteptări

Trecerea în revistă de mai sus a subliniat faptul că indivizii folosesc diferite strategii pentru a evalua veridicitatea unei afirmații, de exemplu credibilitatea acesteia (Dunwoody, 2009; Reder, 1982). Cercetări anterioare asupra distorsiunii negativității indică faptul că informațiile negative pot părea mai credibile (vezi, de exemplu, Hilbig, 2009). Însă, deși valența cadrului este variabila esențială în cercetările asupra distorsiunii negativității, nu este singurul indicator pe care indivizii se pot baza când evaluează credibilitatea și veridicitatea. Mai mult, pot fi conceptualizate interacțiuni între cadrare și alte indicii. De exemplu, Koch și Peter (2017) investighează apariția unei distorsiuni a negativității în contextul comunicării politice și susțin că negativitatea este asociată cu știrile, iar pozitivitatea, cu încercările de persuasiune. În speță, se poate ca indivizii să se fi obișnuit că politicienii transmit știri atunci când oferă informații negative, ceea ce amplifică percepția de adevăr. Prin opoziție, un politician care oferă informații pozitive poate stârni suspiciunea că încearcă să persuadeze receptorii, declanșând

astfel reticență și îndoială, ceea ce duce la diminuarea percepției de veridicitate a mesajului. Aceste rezultate sugerează că alți indicatori ai comunicării — aici, contextul comunicării — pot interacționa cu cadrul și pot astfel modera rezultatele în ce privește distorsiunea negativității. În cazul prezentat de Koch și Peter (2017), contextul ar trebui să amplifice biasul negativității observat în general, dat fiind că o încadrare negativă, în comparație cu una pozitivă, are o probabilitate și mai mare de a fi percepută ca adevărată.

Astfel, conform lui Koch și Peter (2017), contextul determină așteptări în privința a ceea ce este mai plauzibil. Extinzând această noțiune generală a așteptărilor, noi susținem că este relevant faptul dacă mesajul în sine corespunde sau contrazice așteptările indivizilor. Ca exemplu, informațiile că 80% dintre căsătorii durează cel puțin zece ani și 20% dintre divorțuri au loc în primii zece ani sunt foarte diferite, deși cadrarea este identică. Procentele prezentate în afirmații pot influența direct credibilitatea și veridicitatea unei afirmații. Indivizii se pot întreba dacă numărul este coerent cu propriile lor experiențe sau cunoștințe, adică dacă acesta se încadrează într-un interval pe care îl consideră plauzibil. Informațiile neverosimile pot părea suspecte și duc astfel la presupunerea unei tentative de inducere în eroare (Bond et al., 1992). La un nivel mai detaliat, cercetările asupra teoriei încălcării așteptărilor (Expectancy Violation Theory) (Burgoon și Jones, 1976) sugerează la rândul lor că direcția încălcării așteptărilor poate fi esențială: o încălcare pozitivă (informația este mai bună decât era de așteptat) duce la o evaluare în mai mică măsură negativă în comparație cu o încălcare negativă (informația este mai rea decât era de așteptat).

Jaffé și Greifeneder (2019) au testat sistematic efectul așteptărilor indivizilor în privința procentelor și au descoperit prin intermediul a patru studii că așteptările indivizilor moderează distorsiunea negativității. Utilizând o nouă serie de itemi, studiile inițiale din această sferă nu au reușit să reproducă distorsiunea negativității. Ca exemplu, o afirmație inițială și cadrată negativ era „61% dintre femeile care vorbesc germană sunt nemulțumite de propria înfățișare“, iar versiunea cadrată pozitiv era „39% dintre femeile care vorbesc germană sunt mulțumite de propria înfățișare“. Participanții au avut o probabilitate mai mare de a considera versiunea cadrată pozitiv ca fiind adevărată, în comparație cu versiunea cadrată negativ. Autorii au explorat

apoi așteptările indivizilor în privința conținutului. În medie, așteptările participanților erau semnificativ diferite de valorile reale prezentate în studiul inițial. În privința afirmației din acest exemplu, participanții credeau că doar 54% dintre femei ar fi nemulțumite de propria înfățișare (cadrare negativă) și 45% ar fi mulțumite de propria înfățișare (cadrare pozitivă). Mai important, evaluarea în ansamblu a celor opt afirmații a scos la iveală printr-o comparație inter-studii un tipar sistematic ca funcție a supra/subestimării și a evaluărilor adevărului: o distorsiune a negativității în evaluarea adevărului era mai probabil să apară dacă un aspect era supraestimat în versiunea cadrată negativ și subestimat în versiunea cadrată pozitiv. Pe baza acestei descoperiri, într-un studiu ulterior am ajustat sistematic numerele din afirmațiile cadrate negativ, astfel încât să fie mai mici decât ar fi fost de așteptat. De exemplu, noua afirmație a fost „41% dintre femeile care vorbesc germană sunt nemulțumite de propria înfățișare“ (cadrare negativă) sau „59% dintre femeile care vorbesc germană sunt mulțumite de propria înfățișare“ (cadrare pozitivă). Odată cu această schimbare, o distorsiune a negativității putea fi creată pentru aceeași serie de itemi în cazul căreia a fost observată o distorsiune a pozitivității în studiul inițial. Într-un pas ulterior, am scăzut și am crescut sistematic numerele din ambele condiții de cadrare, prezentând afirmații cadrate negativ și pozitiv cu numere mai mari sau mai mici decât era de așteptat. Această configurație ne-a permis să investigăm efectul pe care l-a avut subestimarea în comparație cu supraestimarea în contextul cadrării. Deși așteptările nu au avut efecte consecvente în cazul cadrărilor pozitive, supraestimarea (în comparație cu subestimarea) a dus la o probabilitate mai mare de a considera adevărate afirmațiile cadrate negativ.

Aceste rezultate indică faptul că așteptările în privința conținutului joacă un rol esențial în apariția distorsiunii negativității. Utilizând aceeași serie de afirmații, s-a putut găsi o distorsiune a negativității și o distorsiune a pozitivității, în funcție de gradul în care numerele prezentate se îndepărtau de așteptările indivizilor. Motivul constă probabil în faptul că e o veste bună să afle că, în cazul unui aspect negativ, numerele reale sunt mai mici decât se așteptau, așa că indivizii au tendința de a crede că e adevărat. Invers, e o veste rea să afle că, în cazul unui aspect negativ, numerele reale sunt mai mari decât se așteptau, ceea ce poate declanșa apariția preferinței ca această afirmație să fie falsă. Această tendință este consecventă definiției generale a unei distorsionări egosintonice ca „proces cognitiv sau perceptiv modificat de

necesitatea de a menține și de a amplifica stima de sine“ (Forsyth, 2008, p. 429). Deși astfel de perspective optimiste asupra lumii reflectă o distorsionare, credința în vești bune și ignorarea veștilor rele poate fi o tendință utilă, dat fiind că tendințele optimiste pot susține starea de sănătate și confort (Taylor și Brown, 1988).

Implicații ale efectului așteptărilor

Dacă așteptările influențează rezultatul evaluărilor veridicității, cei interesați de diseminarea de fake news și de politica post-adevăr pot profita de asta. În cel mai rău caz, o afirmație falsă este comunicată într-un mod care sabotează credibilitatea, prin distribuirea sistematică de informații negative aflate sub pragul așteptărilor indivizilor. Drept urmare, este recomandabil ca indivizii să fie foarte atenți la detaliile afirmației și la modul în care aceasta are legătură cu propriile așteptări. Sunt surprins că numerele sunt atât de mari sau atât de mici? Numerele au fost alese astfel încât să depășească sau să nu ajungă la nivelul așteptărilor? Cum mă face să mă simt faptul de a citi o astfel de afirmație? Asemenea întrebări pot fi puse pentru a activa o abordare atentă a informației ambivalente: cineva s-ar putea să facă această afirmație să pară bună și, astfel, adevărată.

Sursele negativității

Trebuie să observăm că distorsiunea negativității a fost analizată utilizându-se afirmații cadrare fie negativ, fie pozitiv. Ce anume înseamnă cadrare negativă și pozitivă? Definiția de lucru este descrierea a ceva într-un mod pozitiv versus a descrie ceva într-un mod negativ, fără a se schimba conținutul. De exemplu, una dintre afirmațiile inițiale ale lui Hilbig (2012b) a avut versiunile: „În Africa, 40% dintre oameni câștigă mai puțin de un dolar pe zi“ (cadrare negativă) și „În Africa, 60% dintre oameni câștigă mai mult

de un dolar pe zi“ (p. 46). Un alt exemplu este: „20% dintre germanii sub 17 ani sunt fumători“ (cadrare negativă) și „80% dintre germanii sub 17 ani sunt nefumători“ (cadrare pozitivă și subliniere). Asta ilustrează că există mai multe modalități de a construi un cadru negativ sau pozitiv: pe de o parte, ne putem concentra asupra valenței conceptului, prezentând un concept negativ (venituri mici) versus pozitiv (venituri puțin mai mari). Pe de altă parte, ne putem concentra asupra negației semantice. Un cadru negativ poate fi constituit prin negarea unui concept pozitiv și un cadru pozitiv poate fi obținut prin negarea unui concept negativ (vezi Jaffé și Greifeneder, 2019).

Cercetări anterioare citate în contextul distorsiunii negativității implicate în evaluarea adevărului sunt adesea legate de conceptul valenței informației (e.g. Baumeister et al., 2001). Cu toate acestea, negarea nu este adusă în discuție la fel de des, însă, după cum am descris anterior, permite construirea unui cadru negativ sau pozitiv. Mai mult, motivele pentru care informația negativă ar putea avea un impact mai puternic poate avea legătură cu negarea, întrucât cercetările arată că encodarea activă a afirmațiilor negare presupune procesare controlată (Deutsch, Kordts-Freudinger, Gawronski și Strack, 2009). La fel ca axarea pe cadrul negativ, negarea ar putea duce la o procesare mai detaliată, care, la rândul său, ar crește credibilitatea.

Pentru a înțelege mai bine efectele cadrării este esențial să cunoaștem ce anume determină cadrarea: valența conceptuală a unei afirmații (negativă versus pozitivă) și dacă afirmația este sau nu negată semantic. Adică ambele forme de negativitate trebuie manipulate separat pentru a înțelege ce provoacă efectul distorsiunii negativității. În acest scop, Jaffé și Greifeneder (2020a) au realizat o serie de studii în care au variat sistematic valența și negarea (vezi Tabelul 7.1 pentru un exemplu).

Tabelul 7.1. Exemplu de item cu variație ortogonală a valenței conceptuale și negării

Valență conceptuală	
Pozitivă	Ne

Negare	Prezentă	30% dintre adulții din Germania nu sunt lipsiți de alergii.	70
	Absentă	70% dintre adulții din Germania sunt lipsiți de alergii.	30

Autorii au descoperit că negarea contribuie la distorsiunea negativității, adică afirmațiile negate au în general o probabilitate mai mare de a fi considerate adevărate. De ce joacă negarea un rol atât de important? Se poate specula că integrarea unei negări într-o propoziție ar putea și să semnaleze pricepere și precizie din partea emițătorului. Dacă este folosită o negare, emițătorul trebuie să aibă un anumit nivel de cunoaștere despre subiect — cum altfel ar putea să folosească formulări atât de specifice?

Pe lângă amplificarea specificității, negarea ar putea și să lărgască sfera stărilor posibile. Dacă revenim la exemplul din Tabelul 7.1, este posibil ca lucrurile să nu stea așa sub mai multe aspecte (de exemplu, „nu suferă de alergii“ pentru că nu au deloc alergii, pentru că nu știu că sunt alergici sau pentru că de fapt nu suferă din cauza lor), dar în general lucrurile stau doar sub un singur aspect (de exemplu, „suferă de alergii“ înseamnă că au o alergie și suferă din cauza ei). Dacă, într-adevăr, se poate să existe mai multe aspecte sub care lucrurile să nu stea într-un fel anume decât să stea într-un fel anume, s-ar putea să fie mai rezonabil să ne bazăm pe aceste diferențe când evaluăm adevărul, mai ales dacă nu avem cunoștințe de specialitate. Mai mult, dacă participanții nu aveau cunoștințe anterioare despre subiect și evaluau credibilitatea pur și simplu încercând să-și imagineze scenarii potențiale, s-ar putea să fie mai simplă imaginarea scenariilor referitoare la versiuni negate și mai dificilă imaginarea scenariilor referitoare la versiunea non-negată, dat fiind că aceasta permite o varietate mai restrânsă.

Și alte cercetări din acest volum se referă la efectul negării (vezi capitolul de Marsh și Stanley, volumul de față). Autorii descriu descoperiri recente conform cărora indivizii se descurcă mai bine la corectarea convingerilor eronate dacă inițial au crezut o afirmație negată și apoi li s-a oferit o afirmație non-negată (în comparație cu situația în care mai întâi aflau o informație non-negată, care apoi trebuia corectată prin aflarea unei afirmații negate). Se poate specula că există o disponibilitate cognitivă mai mare de a înlocui o afirmație negată cu o afirmație mai simplă și mai specifică decât invers.

Implicații ale efectului negării

Până acum, datele din cercetarea realizată de Jaffé și Greifeneder (2020a) indică faptul că negarea este încă o piesă importantă a puzzle-ului care reprezintă modul în care indivizii procedează când evaluează adevărul. Din nou, negarea poate fi utilizată într-o manieră manipulatorie în cazul comunicării de fake news și informații post-adevăr. O afirmație falsă poate conține o negare pentru a părea mai credibilă și pentru a oferi sistematic impresia că informația e adevărată. În scopuri defensive, indivizii ar putea să examineze cu atenție dacă negarea le pare plauzibilă. De ce o anumită informație este prezentată într-o formulare care conține o negație? Dacă afirmația ar fi recadrată, astfel încât să nu conțină negarea? Ar suna diferit?

Distanță psihologică

O a treia posibilă variabilă moderatoare a apariției distorsiunii negativității în evaluarea adevărului este reprezentată de atitudinea indivizilor, dat fiind că aceasta influențează modul în care individul integrează și compară indiciile veridicității. Aici ne concentrăm asupra teoriei nivelului de reprezentare (construal level) (Liberian și Trope, 2008; Trope și Liberman, 2003, 2010), ca teorie a atitudinilor, care conceptualizează modul în care indivizii encodează, procesează și integrează informația. Teoria nivelului de reprezentare (Construal Level Theory, CLT) (Liberian și Trope, 2008) presupune că indivizii se gândesc la obiecte sau evenimente pe un continuum care variază de la concret la abstract. Când se gândesc în detaliu la o pădure, indivizii își pot imagina culorile trunchiului și crengilor și ale diferitelor frunze și au astfel o reprezentare foarte concretă a conceptului de pădure. Cealaltă extremă ar fi să se gândească la o pădure în tipare abstracte absolute, cum ar fi o locație recreațională verde dintr-o zonă rurală. Vedem cu ușurință că, în funcție de unde anume pe acest continuum se află reprezentarea individului, sunt luate în considerare aspecte foarte diferite ale informației. La un nivel concret (sau la un nivel reprezentational scăzut), indivizii includ un număr mare de detalii, dar și de informații subordonate. Invers, la un nivel mai abstract (sau la un nivel reprezentational înalt), indivizii includ mai

puține detalii, dar mai multe elemente abstracte și esențiale ale obiectului sau fenomenului ca informații supraordonate. Teoria nivelului de reprezentare presupune că indivizii au reprezentări situate la un nivel mai scăzut dacă obiectele și evenimentele sunt apropiate din punct de vedere psihologic (temporal, spațial, social sau ca probabilitate) și au reprezentări abstracte pentru obiecte și evenimente distante din punct de vedere psihologic.

Teoria nivelului de reprezentare face presupuneri specifice în privința relevanței informațiilor pozitive și negative (Trope și Liberman, 2010). Autorii susțin că argumentele în favoarea a ceva sunt supraordonate argumentelor în defavoarea a ceva, întrucât importanța subiectivă a aspectelor negative depinde de prezența sau absența aspectelor pozitive în mai mare măsură decât reciproca (Eyal, Liberman, Trope și Walther, 2004; Herzog, Hansen și Wänke, 2007). Un exemplu simplu este reprezentat de tratamentele medicale. Pentru a ilustra: dacă un tratament medical pare benefic (are aspecte pozitive), doar în acest caz individul ar lua în considerare și ar discuta potențialele efecte secundare (aspecte negative). Însă, dacă nu există beneficii (nu există niciun aspect pozitiv), potențialele efecte secundare par irelevante și nu sunt luate în considerare. Echivalând aspectele favorabile cu informațiile pozitive și aspectele nefavorabile cu informațiile negative, teoria nivelului de reprezentare permite predicția că, atunci când indivizii folosesc reprezentări la nivel înalt, informațiile pozitive, comparate cu cele negative, sunt procesate mai detaliat, și astfel sunt considerate adevărate. Această concepție este susținută din mai multe perspective. De exemplu, s-a demonstrat că rezultatele sociale pozitive ale unei acțiuni ies mai mult în evidență pe măsură ce distanța socială crește, iar cadrarea mesajelor persuasive în termeni de câștiguri comparate cu pierderi devine mai puțin semnificativă dacă participanții fac judecăți în privința unor entități distante din punct de vedere social decât dacă ar lua în considerare entități aflate în proximitate (Nan, 2007). Se pare și că indivizii evaluează și experiențele afective negative, și cele pozitive ca fiind mai plăcute în cazul în care folosesc reprezentări abstracte, dat fiind că abstractizarea determină amplificarea aspectelor pozitive ale acestor experiențe (Williams, Stein și Galguera, 2014). Se presupune că, deși indivizii sunt preocupați de rezultatele negative și prohibitive din aici-și-acum (cum ar fi dezamăgirea propriei persoane în cazul în care ar pica un examen), odată cu creșterea distanței psihologice, ei au tendința de a se concentra asupra rezultatelor pozitive și

stimulante (cum ar fi obținerea unei note bune; Pennington și Roese, 2003). Rezumând, aspectele pozitive cântăresc mai greu în condiția de abstractizare sau de distanță psihologică, în timp ce aspectele negative intră în scenă când reprezentările sunt mai concrete și distanța este mai mică.

Acesta este cadrul teoretic în care noi (Jaffé și Greifeneder, 2020b) am emis ipoteza că distorsiunea negativității în evaluarea adevărului este deosebit de pronunțată în condițiile proximității psihologice și în cazul în care reprezentările indivizilor sunt concrete, și este mai puțin intensă în condițiile în care distanța psihologică este mai mare și reprezentările indivizilor sunt abstracte. Am investigat această ipoteză într-o primă serie de studii, folosind afirmații cadrate negativ și pozitiv, referitoare la locuri apropiate versus depărtate din punct de vedere spațial de participanți, manipulând astfel distanța psihologică prin distanța spațială. Rezultatele acestor prime studii oferă dovezi preliminare ale atenuării sau inversării distorsiunii negativității determinate de creșterea distanței psihologice. Cu toate acestea, prima serie de studii trebuie urmată de alte teste și replicări, cum ar fi, de exemplu, investigarea altor dimensiuni ale distanței. Mai mult, pare interesantă examinarea diferențelor idiosincrasice ale distanței psihologice. În funcție de conținutul afirmației, percepțiile relevanței și importanței se pot schimba și o temă foarte relevantă poate părea mai apropiată din punct de vedere psihologic decât o alta despre care indivizilor nu le pasă.

Implicații ale efectului distanței psihologice

Se poate să reprezinte un beneficiu faptul ca indivizii să afle că distanța psihologică provoacă un efect asupra modului în care este evaluată informația. Pentru a se elibera de acest efect, indivizii pot încerca să varieze distanța pe care o au față de afirmație. Ce părere ar avea în privința afirmației dacă s-ar referi la un loc îndepărtat, și nu la orașul lor? Sau dacă s-ar referi la trecut, și nu la prezent? Aplicarea acestei proceduri i-ar putea ajuta pe indivizi să depășească distorsiunea negativității când evaluează adevărul, dat fiind că, prin creșterea distanței psihologice, instanțele negative pot cântări mai puțin și pot avea o probabilitate mai mică de a determina creșterea

veridicității percepute și a credibilității.

Mecanismele subiacente distorsiunii negativității în evaluările adevărului

Perspectiva din prezent indică faptul că sunt mai multe lucruri de spus despre efectul pe care distorsiunea negativității îl are în evaluările adevărului. Cercetările inițiale (Hilbig, 2009, 2012b) au arătat că afirmațiile cadrate negativ, în comparație cu cele cadrate pozitiv, au o probabilitate mai mare de a fi considerate adevărate. Acest capitol susține că un cadru negativ s-ar putea să nu reprezinte întotdeauna o condiție suficientă pentru creșterea credibilității și a veridicității. Alte aspecte, cum ar fi așteptările indivizilor, semantica afirmațiilor și atitudinile indivizilor, ar putea cel puțin să modereze efectul valenței pe care îl are cadrul asupra evaluărilor adevărului.

Examinarea variabilelor care ar putea juca un rol de moderare în privința distorsiunii negativității pare recomandabilă, întrucât permite o înțelegere mai profundă a proceselor subiacente. În literatură au fost discutate mai multe explicații ale perspectivei generale de tipul versiunea-negativă-este-mai-bună-decât-cea-pozitivă, cum ar fi faptul că exemplele negative sunt percepute ca având o mai mare valoare informativă (Peeters și Czapinski, 1990), faptul că exemplele negative au o probabilitate mai mare de a fi elaborate cognitiv (Lewicka, 1997; Peeters și Czapinski, 1990) și faptul că instanțele negative sunt detectate mai ușor, dat fiind că sunt mai rare și mai amenințătoare (Dijksterhuis și Aarts, 2003). Și alte cercetări abordează această concepție, arătând că informațiile pozitive seamănă mai mult unele cu altele decât informațiile negative (Alves, Koch și Unkelbach, 2017; Koch, Alves, Krueger și Unkelbach, 2016). Informațiile pozitive sunt așadar mai degrabă redundante, în timp ce informațiile negative sunt diferențiate, indivizii beneficiind astfel dacă le iau în considerare pe acestea din urmă (Alves et al., 2017). În sfârșit, cadrarea negativă și adevărul pot fi percepute ca fiind asociate, pentru că este mai puțin probabil ca alți indivizi să mintă

când aduc vești proaste (Hilbig, 2009; vezi și Koch și Peter, 2017 pentru o perspectivă asupra comunicării politicienilor). Toate acestea sugerează că procesele subiacente sunt definite de o perspectivă cognitivă puternică. Acest lucru este consecvent rezultatelor cercetărilor anterioare, conform cărora indivizii cu bune abilități matematice sunt mai puțin susceptibili la efectele cadrării (Peters, 2012), ceea ce poate însemna că deosebit de importantă pentru examinarea distorsiunilor implicate în evaluarea adevărului este capacitatea cognitivă. Cu toate acestea, cercetarea realizată de Jaffé și Greifeneder (2019) asupra așteptărilor sugerează că și mecanismele motivaționale pot fi importante, iar acestea au fost neglijate până acum în sfera posibilelor procese subiacente. Pentru a recapitula: autorii au observat că faptele negative mai puțin frecvente (și, astfel, eventual mai puțin negative) decât era de așteptat au o probabilitate mai mare de a fi considerate adevărate. Poate pentru că indivizii atribuie greșit adevărului ușurarea pe care o resimt când află că lucrurile stau mai bine decât se așteptau sau poate pentru că pur și simplu își doresc ca lucrurile bune să fie adevărate.

Directii viitoare

Simpla formulare a informației, fie că e vorba despre valență, negare sau distanță, influențează evaluările adevărului, astfel că implicațiile cercetărilor asupra distorsiunii negativității au o probabilitate mare de a fi extinse și relevante pentru toate tipurile de comunicare, indiferent dacă e scrisă sau orală. Implicațiile se referă, de exemplu, la detectarea înșelării în interacțiunile interpersonale (Bond et al., 1992), la evaluarea adevărului mesajelor transmise de politicieni (Koch și Peter, 2017) sau la răspunsurile la sondaje sau la studiile longitudinale. Cercetări ulterioare ar putea contribui prin investigarea modului în care cadrarea informației influențează judecățile indivizilor în diferite contexte. În cercetările noastre, adesea ne concentrăm asupra domeniilor în care considerăm că există un consens în privința a ceea ce e pozitiv (e.g. sănătate) și a ceea ce e negativ (e.g. boala). Cu toate acestea, analizarea mai multor contexte ar permite examinarea modului în care

convingerile și opiniile idiosincrasice pot interacționa cu efectele cadrării în privința evaluărilor adevărului. Mai mult, studiile ar putea examina și modul în care efectele se pot schimba nu numai în funcție de context, ci și de amplificarea sau reducerea relevanței unor afirmații. De exemplu, în cazul unei afirmații care informează indivizii în privința unui risc personal (în comparație cu un risc care ar afecta un inamic), atributele afirmației s-ar putea să fie procesate diferit.

Mai mult, cercetări ulterioare pot investiga și diferențele potențiale dintre informațiile distribuite offline și cele din mediul online. Dată fiind viteza incredibilă a răspândirii informațiilor pe care o are comunicarea prin internet, sau, mai specific, prin social media, este important să înțelegem diferențele de procesare a informației (vezi Baumeister et al., 2001) și impactul său asupra evaluărilor ulterioare ale adevărului. Se poate specula că indivizii sunt mai predispuși să proceseze în profunzime un articol de ziar pe care îl citesc după micul dejun într-o dimineață de duminică decât un fragment de știre din feed-ul Facebook. Această diferență ar putea modera potențialele distorsiuni ale cadrării. Dacă, de exemplu, negarea este într-adevăr un indicator al competenței emițătorului, conform ipotezei noastre, se poate să aibă un efect mai mare în condiții de intensitate redusă a procesării, cum ar fi feed-ul Facebook (cf. Elaboration Likelihood Model of Persuasion, Petty și Cacioppo, 1986).

Progresul tehnologic a făcut mai simplă distribuirea informației, dar a devenit mai ușoară și verificarea informației. Cu toate acestea, dat fiind volumul masiv de informații distribuite, s-ar putea să se fi redus probabilitatea de implicare în verificarea faptelor. Acest lucru lasă cale liberă efectelor cadrării, negării sau așteptărilor, după cum am arătat aici.

Am susținut pe parcursul acestui capitol că o mai bună cunoaștere a distorsiunii negativității ar putea veni în ajutorul celor care doresc să disemineze fake news și al celor care doresc să se protejeze pe ei înșiși sau să-i protejeze pe alții de manipulare. În lupta pentru adevăr (sau pentru fals), cunoașterea puterii pe care o are modul de exprimare este probabil importantă pentru toți cei implicați.

Concluzie: noi perspective asupra distorsiunii negativității în evaluarea adevărului

În acest capitol am discutat distorsiunea negativității în privința evaluărilor adevărului. Am subliniat motivul pentru care, într-o epocă a post-adevărului, este important să cunoaștem cum procedează indivizii când evaluează adevărul. Analizând în profunzime distorsiunea negativității, am oferit idei noi în ceea ce privește potențiale variabile moderatoare ale distorsiunii negativității și care influențează tendința indivizilor de a crede că o anumită afirmație este adevărată. Am concluzionat că, în special pentru cei care încearcă să anuleze această distorsiune, este important să cunoaștem maleabilitatea evaluărilor adevărului și să știm care sunt modurile în care ajustări minore ale cadrării sau conținutului influențează evaluarea informației.

Referințe

Alves, H., Koch, A. și Unkelbach, C. (2017). Why good is more alike than bad: Processing implications. *Trends in Cognitive Sciences*, 21, 69–79.
<https://doi.org/10.1016/j.tics.2016.12.006>

Anolli, L., Balconi, M., și Ciceri, R. (2003). Linguistic styles in deceptive communication: Dubitative ambiguity and elliptic eluding in packaged lies. *Social Behavior and Personality: An International Journal*, 31, 687–710.
<https://doi.org/10.2224/sbp.2003.31.7.687>

Baumeister, R.F., Bratslavsky, E., Finkenauer, C. și Vohs, K.D. (2001). Bad

is stronger than good. *Review of General Psychology*, 5, 323–370.
<https://doi.org/10.1037//1089-2680.5.4.323>

Bond, C.F., Jr., Omar, A., Pitre, U., Lashley, B.R., Skaggs, L.M. și Kirk, C.T. (1992). Fishy-looking liars: Deception judgment from expectancy violation. *Journal of Personality and Social Psychology*, 63, 969–977.
<http://dx.doi.org/10.1037//0022-3514.63.6.969>

Buller, D.B., Burgoon, J.K., White, C.H. și Ebescu, A.S. (1994). Interpersonal deception VII behavioral profiles of falsification, equivocation, and concealment. *Journal of Language and Social Psychology*, 13, 366–395.
<https://doi.org/10.1177/0261927X94134002>

Burgoon, J.K., Blair, J.P. și Strom, R.E. (2008). Cognitive biases and nonverbal cue availability in detecting deception. *Human Communication Research*, 34, 572–599. <https://doi.org/10.1111/j.1468-2958.2008.00333.x>

Burgoon, J.K. și Jones, S.B. (1976). Toward a theory of personal space expectations and their violations. *Human Communication Research*, 2, 131–146. <https://doi.org/10.1111/j.1468-2958.1976.tb00706.x>

Dechêne, A., Stahl, C., Hansen, J. și Wanke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, 14, 238–257.
<https://doi.org/10.1177/1088868309352251>

Deutsch, R., Kordts-Freudinger, R., Gawronski, B. și Strack, F. (2009). Fast and fragile: A new look at the automaticity of negation processing. *Experimental Psychology*, 56, 434–446. <https://doi.org/10.1027/1618-3169.56.6.434>

Dijksterhuis, A. și Aarts, H. (2003). On wildebeests and humans: The preferential detection of negative stimuli. *Psychological Science*, 14, 14–18.
<https://doi.org/10.1111/1467-9280.t01-1-01412>

Dunwoody, P.T. (2009). Theories of truth as assessment criteria in judgment and decision making. *Judgement and Decision Making*, 4, 116–125.

Eyal, T., Liberman, N., Trope, Y. și Walther, E. (2004). The pros and cons of temporally near and distant action. *Journal of Personality and Social Psychology*, 86, 781–795. <https://doi.org/10.1037/0022-3514.86.6.781>

Forsyth, D.R. (2008). Self-serving bias. În W.A. Darity (coord.), *International encyclopedia of the social sciences* (ed. a II-a, vol. 7, p. 429). Accesat la <http://scholarship.richmond.edu/jepson-faculty-publications>

Grice, H.-P. (1975). Logic and conversation. În P. Cole și J.L. Morgan (coord.), *Syntax and semantics 3: Speech acts* (pp. 41–58). New York: Academic Press.

Hammond, K.R. (1996). *Human judgment and social policy*. New York: Oxford University Press.

Hammond, K.R. (2007). *Beyond rationality*. Oxford: Oxford University Press.

Herzog, S.M., Hansen, J. și Wänke, M. (2007). Temporal distance and ease of retrieval. *Journal of Experimental Social Psychology*, 43, 483–488. <https://doi.org/10.1016/j.jesp.2006.05.008>

Hilbig, B.E. (2009). Sad, thus true: Negativity bias in judgments of truth. *Journal of Experimental Social Psychology*, 45, 983–986. <https://doi.org/10.1016/j.jesp.2009.04.012>

Hilbig, B.E. (2012a). Good things don't come easy (to mind): Explaining framing effects in judgments of truth. *Experimental Psychology*, 59, 38–46. <https://doi.org/10.1027/1618-3169/a000124>

Hilbig, B.E. (2012b). How framing statistical statements affects subjective veracity: Validation and application of a multinomial model for judgments of truth. *Cognition*, 125, 37–48. <https://doi.org/10.1016/j.cognition.2012.06.009>

Jackson, J. (2017). Fact-checkers are weapons in the post-truth wars, but they're not all on one side. *The Guardian*. Accesat la www.theguardian.com/media/2017/feb/15/fact-checkers-are-weapons-in-the-post-truth-wars-but-theyre-not-all-on-one-side

Jaffé, M.E. și Greifeneder, R. (2019). Less than I expected and oh so true? On the interplay between expectations and framing effects in judgments of truth. *Journal of Language and Social Psychology*, 1–21. <https://doi.org/10.1177/0261927X19869392>

Jaffé, M.E. și Greifeneder, R. (2020a). Negative or negated, thus true? An investigation of concept valence and semantic negation as drivers of framing effects in judgments of truth. *Manuscris în curs de pregătire*.

Jaffé, M.E. și Greifeneder, R. (2020b). Negative is true here and now, but not so much there and then: On the impact of psychological distance on the negativity bias. *Manuscris în curs de pregătire*.

Koch, A., Alves, H., Krueger, T. și Unkelbach, C. (2016). A general valence asymmetry in similarity: Good is more alike than bad. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42, 1171–1192. <https://doi.org/10.1037/xlm0000243>

Koch, T. și Peter, C. (2017). Effects of equivalence framing on the perceived truth of political messages and the trustworthiness of politicians. *Public Opinion Quarterly*, 81, 847–865. <https://doi.org/10.1093/poq/nfx019>

Lewicka, M. (1997). Is hate wiser than love? Cognitive and emotional utilities in decision making. În R. Ranyard, W.R. Crozier și O. Svenson (coord.), *Decision making: Cognitive models and explanations*. Londra, Marea Britanie: Routledge.

Lewicka, M., Czapinski, J. și Peeters, G. (1992). Positive-negative asymmetry or „when the heart needs a reason“. *European Journal of Social Psychology*, 22, 425–434. <https://doi.org/10.1002/ejsp.2420220502>

Liberman, N. și Trope, Y. (2008). The psychology of transcending the here and now. *Science*, 322, 1201–1205. <https://doi.org/10.1126/science.1161958>

Marsh, E.J., și Stanley, M.L. (2021). False beliefs: Byproducts of an adaptive knowledge base? În R. Greifeneder, M.E. Jaffé, E.J. Newman, și N. Schwarz (coord.), *Psihologia fake news: acceptarea, distribuirea și corectarea dezinformării* (pp. 131–146). București: Trei.

Nan, X. (2007). Social distance, framing, and judgment: A construal level perspective. *Human Communication Research*, 33, 489–514.
<https://doi.org/10.1111/j.1468-2958.2007.00309.x>

Norman, M. (8 noiembrie 2016). Whoever wins the US presidential election, we've entered a post-truth world: There's no going back now. *The Independent*. Accesat la www.independent.co.uk/voices/us-election-2016-donald-trump-hillary-clinton-whowins-post-truth-world-no-going-back-a7404826.html

Oxford Dictionary. (2016). Oxford Dictionary Word of the Year 2016.

Peeters, G. și Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, 1, 33–60.
<https://doi.org/10.1080/14792779108401856>

Pennington, G.L. și Roese, N.J. (2003). Regulatory focus and temporal distance. *Journal of Experimental Social Psychology*, 39, 563–576.
[https://doi.org/10.1016/S0022-1031\(03\)00058-1](https://doi.org/10.1016/S0022-1031(03)00058-1)

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science*, 21, 31–35. <https://doi.org/10.1177/0963721411429960>

Petty, R.E. și Brinol, P. (2008). Persuasion: From single to multiple to metacognitive processes. *Perspectives on Psychological Science*, 3, 137–147.

Petty, R.E. și Cacioppo, J.T. (1986). The elaboration likelihood model of persuasion. În L. Berkowitz (coord.), *Advances in experimental social psychology* (vol. 19, pp. 123–205). [https://doi.org/10.1016/S0065-2601\(08\)60214-2](https://doi.org/10.1016/S0065-2601(08)60214-2)

Pratto, F. și John, O.P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, 61, 380–391. <https://doi.org/10.1037/0022-3514.61.3.380>

Reber, R. și Greifeneder, R. (2017). Processing fluency in education: How

metacognitive feelings shape learning, belief formation, and affect. *Educational Psychologist*, 52, 84–103.
<https://doi.org/10.1080/00461520.2016.1258173>

Reber, R. și Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, 8, 338–342.
<https://doi.org/ccog.1999.0386>

Reder, L.M. (1982). Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. *Psychological Review*, 89, 250–280.
<https://doi.org/10.1037//0033-295X.89.3.250>

Reinhard, M.-A., Greifeneder, R. și Scharmach, M. (2013). Unconscious processes improve lie detection. *Journal of Personality and Social Psychology*, 105, 721–739. <https://doi.org/10.1037/a0034352>

Schwarz, N. și Jalbert, M. (2021). When news feels true: Intuitions of truth and the acceptance and correction of misinformation. În R. Greifeneder, M.E. Jaffé, E.J. Newman și N. Schwarz (coord.), *Psihologia fake news: acceptarea, distribuirea și corectarea dezinformării* (pp. 73–89). București, Trei.

Shiv, B., Britton, J.A.E. și Payne, J.W. (2004). Does elaboration increase or decrease the effectiveness of negatively versus positively framed messages? *Journal of Consumer Research*, 31, 199–208. <https://doi.org/10.1086/383435>

Stanley, M.L., Yang, B.W. și Marsh, E.J. (2019). When the unlikely becomes likely: Qualifying language does not influence later truth judgments. *Journal of Applied Research in Memory and Cognition*, 8, 118–129.
<https://doi.org/10.1016/j.jarmac.2018.08.004>

Taylor, S.E. și Brown, J.D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103, 193–210. <https://doi.org/10.1037/0033-2909.103.2.193>

Trope, Y. și Liberman, N. (2003). Temporal construal. *Psychological Review*, 110, 403–421. <https://doi.org/10.1037/0033-295X.110.3.403>

Trope, Y. și Liberman, N. (2010). Construal-level theory of psychological

distance. *Psychological Review*, 117, 440–463.
<https://doi.org/10.1037/a0018963>

Unkelbach, C. și Koch, A. (2019). Gullible but functional? Information repetitions and the formation of beliefs. În J.P. Forgas și R.F. Baumeister (coord.), *The social psychology of gullibility: Conspiracy theories, fake news and irrational beliefs* (pp. 42–60). Abingdon-Thames: Routledge.

Unkelbach, C. și Rom, S.C. (2017). A referential theory of the repetition-induced truth effect. *Cognition*, 160, 110–126.
<https://doi.org/10.1016/j.cognition.2016.12.016>

Vrij, A. și Mann, S. (2001). Who killed my relative? Police officers' ability to detect real-life high-stake lies. *Psychology, Crime & Law*, 7, 119–132.
<https://doi.org/10.1080/10683160108401791>

Williams, L.E., Stein, R. și Galguera, L. (2014). The distinct affective consequences of psychological distance and construal level. *Journal of Consumer Research*, 40, 1123–1138. <https://doi.org/10.1086/674212>.

Capitolul 8. Convingeri eronate

Efecte secundare ale unei baze de cunoștințe adaptative?

Elizabeth J. Marsh și Matthew L. Stanley

Pizzagate. Masacrul de la Bowling Green*****. Susținerea acordată lui Donald Trump de către Papa Francisc. Asemenea „fake news“ ajung adesea virale, ocolind indiciile tipice ale credibilității și inventând lucruri despre persoane și locuri familiare. Este aproape imposibilă eliminarea completă a expunerii la informații incorecte (dezinformare), mai ales pentru că acestea tind să nu dispară (dacă dispar vreodată), fiind menținute prin internet (Shin, Jian, Driscoll și Bar, 2018). Deși unele nu au avut niciodată pretenția de a fi crezute (e.g. satirele din The Onion), altele sunt îngrijorătoare pentru că acționează ca instrumente active de propagandă (e.g. știrile atribuite boților ruși***** de pe Twitter și Facebook). Adesea, acestea pun accentul pe emoții și motivații prin frica provocată cititorilor, de exemplu folosind imagini cu imigranți care traversează granița de sud sau prin informații false despre viitoarea implementare a unor „comisii birocratice de condamnare la moarte“, care au puterea de a opri intervenția medicală acordată indivizilor.

Dar această „gândire la cald“ apare într-un sistem menit să învețe, să mențină și să actualizeze cunoștințele și convingerile despre lume. Probabilitatea de a ajunge să credem fake news (și capacitatea noastră de a corecta convingerile eronate) depinde de procesele cognitive de zi cu zi, subiacente cunoașterii și convingerilor, și nu de procese sau strategii unice sau speciale (Isberner și

Richter, 2014; Marsh, Cantor și Brashier, 2016; Rapp și Donovan, 2017; Rapp și Salovich, 2018). Începem acest capitol cu descrierea succintă a bazei de cunoștințe, definind conținuturile acesteia și comparând-o cu alte forme de memorie. Apoi descriem cinci proprietăți ale acesteia care adesea o fac eficientă și puternică — dar uneori duc la convingeri eronate, cu implicații pentru motivele care îi fac pe oameni să creadă știri false. În ultima secțiune a capitolului rezumăm strategiile care se pot dovedi valoroase pentru contracararea problemei fake news.

Baza de cunoștințe

Oamenii știu multe lucruri și aceste cunoștințe pot lua multe forme, de la fapte concrete (e.g. care e capitala statului Peru) la scenarii (abstractizări ale experiențelor anterioare, cum ar fi scenariul de navigare prin protocoalele de securitate ale aeroportului) la înțelegerea sistemelor complexe (e.g. cum funcționează economia și cum fluctuația dobânzilor are implicații pentru alte părți ale sistemului). Este acumulată pe parcursul întregii vieți, uneori formal (la școală, prin instruire intenționată), alteori, informal (prin încercare-și-eroare, expunere la media, interacțiuni interpersonale etc.).

În științele psihologice, cunoașterea este adesea definită în funcție de ce nu e: nu înseamnă să ne gândim la un eveniment specific care a avut loc într-un anumit loc și într-un anumit moment din trecut. Diferența dintre cunoaștere și actualizarea anumitor evenimente (memorie episodică) este poate cel mai bine ilustrată prin exemple. Se poate ca o persoană să știe ce e un bal de absolvire, de exemplu, dar își poate aminti un episod anume care implică primirea unui buchet roz pe care l-a purtat la propria festivitate. Se poate ca un individ să știe care sunt pașii pe care trebuie să-i urmeze pentru o călătorie cu avionul sau se poate să se gândească la un zbor deosebit de turbulent din experiența sa, de la Columbus la St. Louis. Se poate ca un individ să știe cine a fost Abraham Lincoln fără să-și amintească să fi învățat despre el la o anumită lecție de la un curs anume.

Noi sugerăm că lupta împotriva erorilor din sistemul de cunoștințe (convingeri false, iluzia cunoașterii, concepții eronate) poate necesita alte strategii decât cele recomandate uneori pentru combaterea amintirilor false ale unor evenimente specifice. Corectarea convingerilor eronate pe care o persoană le are în privința legăturii între vaccinuri și autism, de exemplu, e diferită de ajutarea unui individ să-și amintească acurat și complet un anumit eveniment. Unele dintre sugestiile din prezent pentru combaterea fake news se bazează pe literatura amplă asupra modului în care oamenii își amintesc greșit evenimente specifice din trecutul lor (e.g. strategii de a se gândi la o anumită persoană, sau sursă, care a prezentat informația într-un anumit moment și într-un anumit loc și dacă acea persoană este demnă de încredere sau are o competență adecvată; Lewandowsky, Ecker și Cook, 2017; Rapp și Salovich, 2018). Se poate ca aceste strategii să funcționeze în unele contexte, însă noi sugerăm că alte strategii ar putea fi mai adecvate pentru corectarea convingerilor eronate implementate de fake news.¹

Să ne gândim la convingerea împărtășită de oamenii care au ajuns să considere în mod eronat că Barack Obama s-a născut în afara SUA. Desigur, e posibil ca oamenii să aibă amintiri false despre un eveniment specific în care au auzit afirmația că Barack Obama s-a născut (sau nu) în afara SUA. Cu toate acestea, noi sugerăm că indiferent dacă oamenii își pot aminti acurat și complet evenimentul în care au fost expuși la această informație e mai puțin important decât dacă au în prezent convingerea că afirmația „Barack Obama s-a născut în afara SUA” e adevărată. Convingerea eronată trebuie corectată, nu amintirea de a o fi auzit. Ceea ce sugerează că trebuie să căutăm mai degrabă în literatura asupra corectării convingerilor, și nu în literatura despre cum pot fi ajutate persoanele să-și amintească acurat evenimente specifice din trecut.

Pe scurt, în acest capitol descriem succint mai multe proprietăți ale bazei de cunoștințe cu implicații în privința procesării fake news, după care vom discuta implicațiile în ceea ce privește combaterea acestei probleme.

Principiul nr. 1

Avem tendința să credem că informațiile sunt adevărate. În general, când aflăm informații noi, oamenii au tendința de a presupune mai degrabă că sunt adevărate, nu false. Gilbert (1991) susține că această distorsiune este eficientă din punct de vedere cognitiv și înseamnă că nu trebuie decât să etichetăm informațiile dacă sunt false (altfel putem presupune că sunt adevărate). Se presupune că această distorsiune reflectă o dispoziție de a crede într-o lume care, în mare, e sinceră, în care e mai probabil ca informațiile pe care le aflăm să fie adevărate. Dar, conform acestui model, etichetarea unei informații ca fiind falsă presupune efort și implică un al doilea pas, incredulitatea (unbelieving), ceea ce reprezintă o problemă dacă acest proces este perturbat.

Mai multe studii susțin această predicție, și anume că perturbările care au loc în timp ce aflăm informații false ne fac să ni le amintim mai târziu ca fiind adevărate (e.g. Gilbert, Tafarodi și Malone, 1993). În cadrul unui studiu (Gilbert, Krull și Malone, 1990), participanții au simulat învățarea unei limbi străine într-un cadru de imersiune în care li se cerea să ghicească semnificația cuvintelor înainte de a primi feedback. Participanții au citit o serie de traduceri din Hopi***** în engleză, fiecare etichetată apoi ca fiind adevărată sau falsă; un aspect esențial a fost faptul că unele serii au fost întrerupte de un semnal sonor la care participanții trebuiau să reacționeze prin apăsarea unei taste. Mai târziu, participanții au completat un test de tip adevărat/fals. Conform predicției, s-a întâmplat mai frecvent ca seriile false întrerupte să fie mai târziu amintite în mod greșit ca fiind adevărate, fapt consecvent ideii că semnalul sonor a interferat cu procesul adițional, care presupune efort, de scepticism la adresa traducerilor false.

Un al doilea exemplu: în cadrul unui alt experiment, Gilbert și colegii săi (1990) le-au cerut participanților să se uite la o înregistrare și apoi să evalueze dacă zâmbetele erau sincere sau prefăcute. După ce vedeau fiecare față, participanților li se spunea explicit dacă zâmbetul era sincer sau prefăcut. În unele serii, efortul cognitiv era amplificat prin solicitarea participanților de a răspunde pe cât de repede posibil la un semnal sonor sau de a spune care din două semnale sonore avea o tonalitate mai înaltă. În faza ulterioară de testare, participanții au avut o probabilitate mai mare de a face erori și de a susține că persoanele din înregistrări zâmbiseră sincer dacă

eticheta sincer/prefăcut a fost asociată cu sarcina secundară. La fel ca în experimentul cu traduceri din Hopi în engleză, explicația a fost că un mai mare efort cognitiv i-a făcut pe participanți să sară peste etapa de scepticism, care implică efort. În consecință, evaluările oferite la teste au reflectat doar stadiul inițial de credință-consimțământ (vezi și Mandelbaum, 2014).

Dacă Gilbert și colegii săi au dreptate, în general oamenii au tendința de a crede titluri precum „Clint Eastwood refuză să accepte Medalia Libertății oferită de Obama, motivând că «el nu e președintele meu»“ și să devină sceptici doar după un al doilea pas cognitiv, care presupune efort.² Ceea ce implică faptul că oamenii distrași sau ocupați au o probabilitate mai mică de a fi sceptici față de informațiile false. Dacă dăm crezare perspectivei lui Gilbert și a colegilor săi, oamenii caracterizați de „lene cognitivă“ (i.e. care, când află informații noi, nu se angajează în procese de reflecție care presupun efort) au o probabilitate mai mare de a crede titluri de fake news relativ incredibile, în timp ce persoanele care tind să aplice procese de reflecție care presupun efort au o probabilitate mai mică de a crede aceleași titluri neverosimile de fake news (Bronstein, Pennycook, Bear, Rand și Cannon, 2019; Pennycook și Rand, 2019). Cu alte cuvinte, persoanele care au oferit răspunsul evident, dar greșit, la întrebări precum „Câți metri cubi de pământ sunt într-o groapă adâncă de trei metri, cu o lungime și o lățime de trei metri?“ au avut de asemenea o probabilitate mai mare de a accepta titluri de fake news (notă: nu e deloc pământ în groapă). La fel, indivizii cu o probabilitate mai mare de a folosi procese reflexive, care implică efort, când află informații noi au o probabilitate mai mare de a refuza teoriile conspirației în comparație cu cei care nu aplică aceleași procese (Swami, Voracek, Stieger, Tran și Furnham, 2014).

Trebuie să facem o remarcă în legătură cu principiul nr. 1: acest efect apare când oamenii sunt obligați să aleagă între adevărat și fals. Dacă nehotărârea este o opțiune, acest efect dispare (Street și Richardson, 2015; vezi și Street și Kingstone, 2017). Cu alte cuvinte, principiul nr.1 descrie o euristică utilizată într-o situație specifică. Dacă trebuie făcută o evaluare de tip adevărat/fals, această euristică se aplică pentru că oamenii presupun că majoritatea celorlalți spun adevărul în cea mai mare parte a timpului (DePaulo, Kashy, Kirkendol, Wyer și Epstein, 1996).

Principiul nr. 2: ne folosim de scurtături pentru a evalua adevărul

În general, oamenii folosesc euristici pentru a evalua dacă informațiile pe care le află sunt adevărate, considerând informația ușor de procesat ca având o probabilitate mai mare de a fi adevărată decât informația mai dificil de procesat. Adică afirmațiile ușor de procesat (de citit, de auzit, de înțeles) au o probabilitate mai mare de a fi considerate adevărate în comparație cu afirmații mai dificil de procesat. În consecință, afirmațiile tipărite cu un font ușor de citit au o probabilitate mai mare de a fi considerate adevărate decât cele tipărite cu un font mai dificil de citit (Reber și Schwarz, 1999). La fel, aforismele cu rimă (ai carte, ai parte) sunt considerate ca având o probabilitate mai mare de a fi adevărate decât cele care nu rimează („banu-i ochiul dracului”; McGlone și Tofighbakhsh, 2000).

Repetiția este poate modalitatea cea mai simplă de a face unele afirmații să pară mai ușor de procesat decât altele, dat fiind că afirmația va fi citită mai ușor (mai repede) a doua oară. Experimental, în general asta se demonstrează cerându-li-se participanților ca într-o primă etapă să citească o serie de informații de cultură generală și apoi, într-o etapă ulterioară, să evalueze adevărul unei serii mai mari de afirmații (de obicei, o jumătate formată din propoziții din prima serie și o jumătate formată din afirmații noi). Afirmațiile repetate sunt considerate ca având o probabilitate mai mare de a fi adevărate decât noile afirmații, fenomen numit efectul iluziei de adevăr. Demonstrat inițial în anii 1970 (Hasher, Goldstein și Toppino, 1977), acest efect a fost reprodus de multe ori, se menține indiferent de variațiile procedurale și de materiale și poate dura multe săptămâni (vezi Dechêne, Stahl, Hansen și Wänke, 2010 pentru o meta-analiză).

Efectul iluziei de adevăr nu este doar un fenomen de laborator și nici nu este limitat la afirmații obscure de cultură generală: a fost demonstrat și în cazul titlurilor de fake news văzute pe Facebook (Pennycook, Cannon și Rand,

2018). Pennycook et al. (2018) au arătat că o unică expunere scurtă la un titlu fake news (e.g. Donald Trump și-a trimis propriul avion pentru a transporta 200 de infanteriști izolați“) a determinat creșterea probabilității ca oamenii să creadă că titlul era adevărat. Această amplificare a acurateții percepute a titlurilor fake news era prezentă după o săptămână, iar expuneri suplimentare au amplificat încrederea în veridicitatea titlurilor. În mod surprinzător, aceleași efecte au fost obținute chiar și când participanților li s-a spus că articolele erau puse la îndoială de un grup independent care se ocupa de verificarea faptelor. Aceste descoperiri sugerează că etichetarea informației ca fiind falsă sau îndoielnică (după cum procedează organizații care se ocupă de verificarea faptelor, ca PolitiFact și Snopes) s-ar putea să nu reprezinte cea mai eficientă modalitate de combatere a fake news și a dezinformării, aspect la care vom reveni mai târziu în acest capitol.

Principiul nr. 3: interpretăm informațiile noi

Fake news reprezintă un fenomen interesant pentru că se referă în general la oameni și locuri reale, ca Hillary Rodham Clinton și Rusia; rareori se referă la oameni sau locuri despre care nu am mai auzit până atunci. Actorii și spațiile familiare înseamnă că ne bazăm pe cunoștințele anterioare când procesăm fake news, ceea ce le face ușor de înțeles și de elaborat. Mai mult, avem un scenariu în privința știrilor — o serie de așteptări despre tipurile de informații pe care le conțin de obicei: cine, ce, când, unde, de ce, precum și povești de viață. Avem oarece idee la ce să ne așteptăm când citim acest tip de știre.

Dat fiind că obiectivul nostru este în general extragerea de sens și integrarea noilor informații în schemele și structurile noastre de convingeri, nu stocăm informația fără să o prelucrăm, și rareori ne amintim formularea exactă a conversațiilor sau a altor forme de comunicare. De exemplu, propoziția „expertul în karate a lovit cărămida“ e amintită greșit ca „expertul în karate a

spart cărămida“ (Brewer, 1977). Un efect similar apare în domeniul vizual; oamenii își amintesc în mod fals că au văzut acțiuni cauzale (e.g. un picior lovind o minge) care erau doar implicite în filmarea video (Strickland și Keil, 2011).

Mai mult, avem tendința inerentă de a trece dincolo de ce e dat, de a face inferențe și de a acoperi golurile. Cu alte cuvinte, sistemul este eficient din punct de vedere cognitiv, bazându-se pe trecut pentru a se pregăti pentru viitor. Asimilăm informații care nu sunt expuse explicit. De exemplu, copiilor mici nu e nevoie să li se spună explicit că „delfinii comunică prin țipete ascuțite“ după ce află că „delfinii călătoresc în grup“ și că „grupurile comunică prin țipete ascuțite“ (Bauer și San Souci, 2010). Transferăm cunoștințele în noi situații și în noi cadre. De exemplu, identificăm autorul unui tablou pe care îl vedem pentru prima oară, pe baza experiențelor noastre anterioare cu opera aceluia artist (Kornell și Bjork, 2008).

Capacitatea noastră de interpretare ne permite să ne descurcăm cu inputuri imperfecte, atâta timp cât „seamănă suficient de bine“. Este bine cunoscută lipsa de fluentă a discursului — ne bâlbâim, facem greșeli de vorbire și folosim referenți ambigui. Iar sistemul este configurat astfel încât să se descurce cu inputuri imperfecte, acceptând un anumit grad de „deviere“ a mesajului (Ferreira, Bailey și Ferraro, 2002). Un cuvânt este recunoscut în multe fonturi, accente și contexte diferite — o astfel de flexibilitate reprezintă un avantaj când avem de-a face cu variații noi ale informației stocate. Pe de altă parte, un sistem flexibil are efectul secundar de a rata înexactitățile care „seamănă suficient de bine“. De exemplu, oamenii adesea nu identifică informațiile incorecte din întrebări, chiar și când sunt avertizați cu privire la asta. Multe persoane vor încerca să ofere un răspuns numeric la întrebarea „Câte animale din fiecare specie a luat Moise pe arcă?“, chiar dacă sunt instruiți în privința acestui truc și mai târziu demonstrează că știu că întrebarea ar fi trebuit să se refere la Noe, nu la Moise („iluzia Moise“; Bottoms, Eslick și Marsh, 2010; Erickson și Mattson, 1981).

Multe strategii diferite au fost implementate în încercarea de a reduce sau de a elimina „iluzia Moise“ și alte fenomene asemănătoare. Însă, ironic, multe dintre aceste strategii au avut efecte inverse, sporind probabilitatea ca oamenii să-și amintească sau să se bazeze pe informații false. De exemplu,

Eslick, Fazio și Marsh (2011) au încercat să atragă atenția participanților asupra informațiilor false din afirmații prin scrierea acestora cu font roșu. Dar asta nu a făcut decât să crească probabilitatea ca participanții să-și amintească aceste erori la un test ulterior. Însă o strategie care chiar pare să reducă influența informațiilor greșite este încurajarea indivizilor să joace rolul unei persoane însărcinate să verifice faptele atunci când citește informații noi (Rapp, Hinze, Kohlhepp și Ryskin, 2014; Brashier, Eliseev și Marsh, 2020). Această abordare a amplificat probabilitatea ca participanții să identifice informațiile false și a redus probabilitatea de a se baza pe acestea mai târziu.

Principiul nr. 4: cunoștințele sunt resimțite diferit

Fenomenologia actualizării cunoștințelor este foarte diferită de ceea ce resimte persoana când reactualizează o amintire (ceea ce implică întoarcerea mentală în timp și spațiu la un anumit eveniment). Cunoștințele „apar în minte” și sunt etichetate ca „știute”, și nu „amintite” (Tulving, 1972). Când actualizează faptul că George Washington a fost primul președinte al SUA, de exemplu, majoritatea oamenilor nu se gândesc la o situație specifică în care au aflat această informație, ci pur și simplu o știu. Desigur, pot exista excepții, persoana amintindu-și circumstanțele în care a aflat informația (e.g. Conway, Gardiner, Perfect, Anderson și Cohen, 1997), dar, în majoritatea cazurilor, lucrurile pe care le știm sunt relativ decontextualizate și separate de situația inițială în care le-am aflat. Deși este mai mică probabilitatea ca oamenii să fie convinși de surse percepute ca nefiind de încredere (Guillory și Geraci, 2013; Mills și Jellison, 1967; Priester și Petty, 1995), în general sursa informației este uitată chiar și după perioade relativ scurte (e.g. Underwood și Pezdek, 1998) și pe măsură ce persoana întâlnește aceeași informație în mai multe contexte diferite (Watkins și Kerkar, 1985).

În prezent, o sugestie populară de abordare a fake news presupune instruirea oamenilor pentru a evalua mai atent și pentru a-și aminti sursele informațiilor

(Rapp și Salovich, 2018; Swire și Ecker, 2018). Acest sfat se bazează pe literatura amplă despre modul în care oamenii își amintesc inexact trecutul, fapt a cărui cauză adesea poate fi găsită în probleme de monitorizare a surselor (Johnson, Hashtroudi și Lindsay, 1993): o presupunere implicită într-o întrebare este atribuită în mod greșit experienței; o afirmație dintr-o expunere este atribuită greșit unei înregistrări video; sau o fotografie oferă detalii încorporate în memoria persoanei. Soluția firească este atragerea atenției asupra originii amintirilor, fie la momentul aflării informațiilor (e.g. ignorarea informațiilor dintr-o sursă cu credibilitate scăzută), fie prin diferențierea surselor în momentul testului (Lindsay și Johnson, 1989). Această abordare este susținută de organizații profesionale ca IFLA (International Federation of Library Associations and Institutions, Federația Internațională a Asociațiilor și Instituțiilor Bibliotecare), care a formulat mai multe recomandări pentru identificarea fake news, inclusiv evaluarea calității website-ului pe care apar și credibilitatea autorului.

Dar monitorizarea surselor este improbabil să rezolve problema reprezentată de fake news (Marsh și Yang, 2017). În afară de faptul că oamenii nu sunt întotdeauna de acord în privința credibilității unei surse (e.g. Fox News este credibilă pentru unii, nu și pentru alții), credibilitatea sursei este adesea inferată pe baza indicatorilor care pot induce în eroare, cum ar fi numărul de citări (Putnam și Phelps, 2017) sau ușurința cu care mesajul poate fi citit (Fragale și Heath, 2004). De fapt, simplitatea procesării informației are în general o probabilitate mai mare de a influența evaluarea adevărului decât sursa informației: afirmațiile care inițial au fost etichetate ca false sau care au fost obținute din surse îndoielnice au totuși o probabilitate mai mare de a fi considerate adevărate în comparație cu afirmațiile noi (Begg și Armour, 1991; Begg, Anas și Farinacci, 1992; Stanley, Yang și Marsh, 2019).

Chiar dacă sursele sunt evaluate corect și informația este stocată în memorie, este probabil să nu aibă niciun efect asupra evaluărilor ulterioare, dat fiind că tendința noastră către eficiența cognitivă înseamnă că avem tendința să folosim informația fără să ne gândim instantaneu de unde o știm (ceea ce reflectă faptul că o mare parte din lucrurile pe care le știm sunt decontextualizate, Tulving, 1972). Rezultatul este că oamenii pot evalua dacă afirmațiile provin dintr-o sursă credibilă sau necredibilă (la două săptămâni după expunere, dacă li se cere explicit), și totuși nu pot să aplice această

informație când evaluează adevărul afirmațiilor (Henkel și Mattson, 2011).

Mai mult, întrucât informațiile sunt asociate mai multor surse, amintirea faptului că o informație anume provine dintr-o sursă cu credibilitate scăzută nu înseamnă neapărat că ar trebui ignorată (Marsh, Meade și Roediger, 2003). Cu alte cuvinte, e posibil ca Donald Trump să fi spus același lucru ca Nancy Pelosi și, astfel, sursa cu credibilitate scăzută este ignorată din cauza amintirii (sau atribuirii greșite) sursei credibile. În al treilea rând, o problemă majoră este că, în timp, sursa informației pur și simplu nu este ținută minte. Aceasta este problema ilustrată în clasicul „efect de adormire” („sleeper effect”) — o sursă cu credibilitate scăzută inițial nu are putere de convingere, dar, în timp, mesajul are un efect mai mare, sursa informației fiind uitată (e.g. Underwood și Pezdek, 1998).

Principiul nr. 5: nu întodeauna ne amintim și/sau aplicăm ce știm

Oamenii consideră adesea cunoștințele ca fiind mereu utilizate dacă sunt stocate în memorie, dar faptul că știm ceva nu reprezintă o garanție că ne vom aminti sau că vom aplica acel lucru când va fi cazul. Cel mai frecvent exemplu al acestui fenomen este senzația „pe-vârful-limbii” (vezi Brown, 1991). Aceasta apare cel mai adesea când încercăm să ne amintim termeni specifici sau nume și presupune senzația frustrantă că știm ceva „ce ne scapă”. În această stare este prezentă hiperconștientizarea problemei, însă alteori se poate ca persoana să nu fie conștientă de faptul că nu poate reactualiza o informație. În general, informația care nu a fost utilizată recent are o probabilitate mai mică de a fi reactualizată spontan (Cantor, Eslick, Marsh, Bjork și Bjork, 2015).

Am oferit deja un exemplu al principiului general conform căruia cunoștințele nu sunt întotdeauna utilizate și am făcut observația că sistemul acceptă inputuri care „seamănă suficient de bine” chiar și dacă știm cum stau

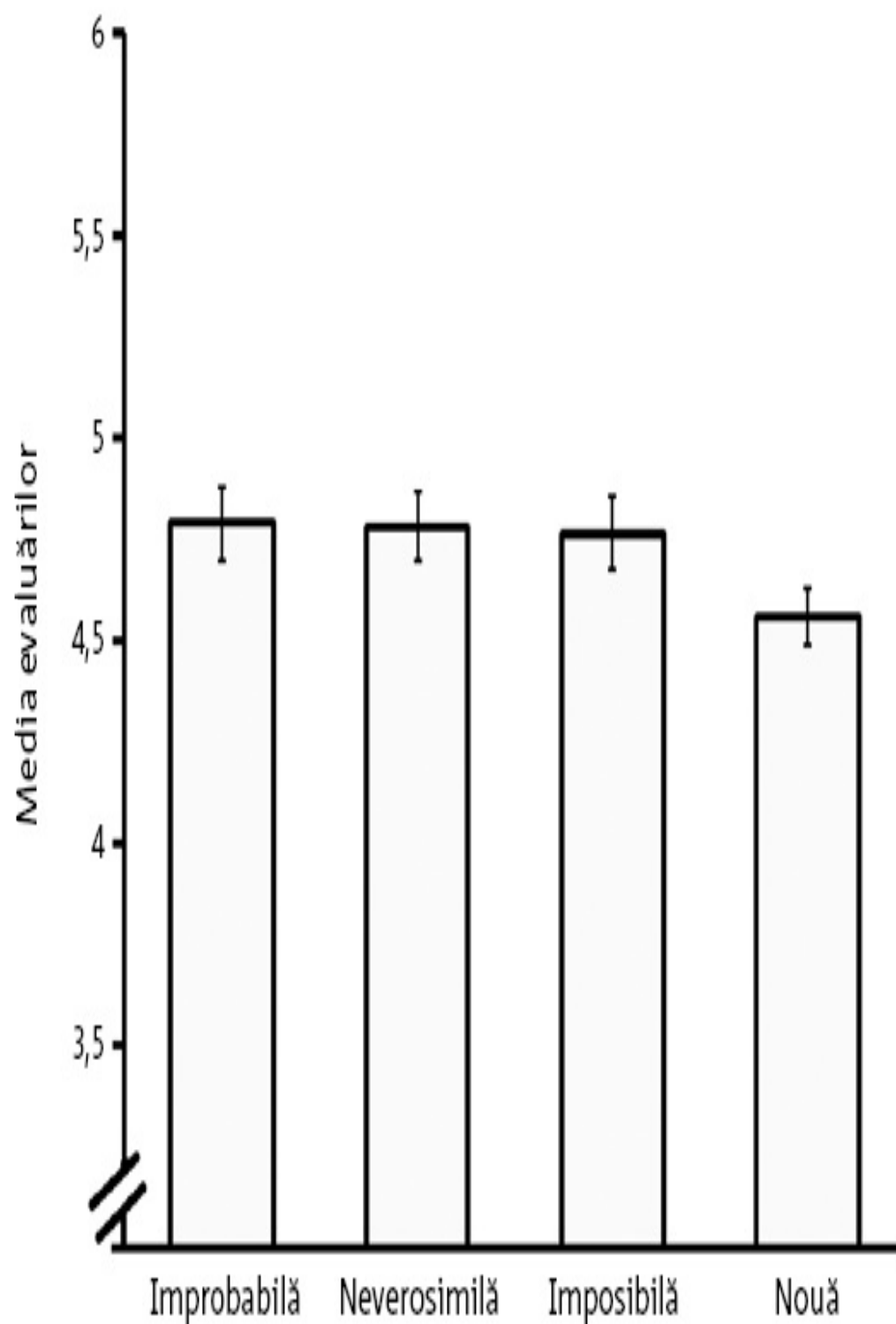
lucrurile de fapt (e.g. nereușita de a sesiza eroarea din întrebarea „Câte animale din fiecare specie a luat Moise pe arcă?“). Această problemă apare chiar și în cazul informațiilor bine știute. Cantor și Marsh (2017) au formulat întrebări de biologie (e.g. „Apa conține doi atomi de heliu și câți atomi de oxigen?“) și de istorie care conțineau presupoziii incorecte („Cum se numeau taxele pe care kongresul britanic le-a impus în colonii pentru zahăr, ceai și timbre?“) și pe care le-au adresat doctoranzilor în biologie și istorie. Tuturor participanților li s-a spus să nu răspundă la întrebările care conțineau presupoziii greșite, precum cele ilustrate printr-un exemplu. Acești specialiști în disciplinele pe care le studiau au avut rezultate puțin mai bune în domeniul în care își scriau teza, dar și ei au ratat aproximativ 30% dintre presupunerile greșite din întrebările din domeniul lor (Cantor și Marsh, 2017).

Un al doilea exemplu implică efectul de iluzie a adevărului; în literatura de specialitate, presupunerea a fost că euristica fluenței intră în scenă în cazul ignoranței, când persoana nu se pricepe. Dar dovezi recente sugerează că lucrurile nu stau întotdeauna așa. De exemplu, expunerea la informații evident false (e.g. „Fusta plisată scurtă pe care o poartă scoțienii se numește sari“) determină creșterea probabilității de a fi ulterior considerate adevărate în comparație cu afirmații noi. Acest efect apare indiferent dacă cunoștințele sunt estimate pe baza normelor de cultură generală sau dacă sunt măsurate în cazul unui individ anume (Fazio, Brashier, Payne și Marsh, 2015). Nu e clar cât de departe pot ajunge aceste efecte; rezultatele inițiale sugerează o condiție limită — în cazul afirmațiilor necredibile cum ar fi „Pământul este un pătrat perfect“ (Pennycook et al., 2018), dar cercetări mai recente sugerează că această descoperire este determinată de un efect de podea (floor effect), și nu de natura itemilor (Fazio, Rand și Pennycook, 2019).

Un al treilea exemplu implică încrederea în informații etichetate clar ca fiind false. Într-un studiu realizat de noi (Stanley et al., 2019), participanții au evaluat afirmații referitoare la informații specifice (e.g. „Este improbabil ca trunchiul unui pom putrezit să se numească copac“ sau „Este imposibil ca studiul șerpilor să se numească ofiologie“). În acest experiment esențial, toți determinanții au fost negativi, adică fiecare afirmație în parte era formulată ca imposibilă, improbabilă sau neverosimilă. Două zile mai târziu, participanții au evaluat adevărul acestor afirmații fără determinanți, precum și afirmații

noi. Primele afirmații au fost considerate ca având o probabilitate mai mare de a fi adevărate decât cele noi (vezi Figura 8.1), chiar dacă fiecare dintre determinanții inițiali pune la îndoială veridicitatea fiecărei afirmații. Când li s-a cerut, majoritatea subiecților au identificat corect determinanții prezentați în prima parte a experimentului; știau că afirmațiile făceau parte dintr-o serie cu o probabilitate foarte scăzută de a fi adevărate, și, cu toate acestea, expunerea la această serie a provocat creșterea probabilității de a considera „adevărată” o afirmație cu un determinant negativ.

Figura 8.1. Media evaluărilor afirmațiilor inițial formulate ca improbabile, neverosimile sau imposibile în comparație cu noile afirmații



Implicații pentru problema fake news

Până acum ne-am concentrat asupra procesării fake news, subliniind faptul că nu suntem prea optimiști în privința intervențiilor centrate exclusiv asupra credibilității sursei. Deși este atrăgătoare perspectiva de a lăsa problema în seama altora (e.g. organizațiilor care se ocupă de verificarea faptelor, ca PolitiFact sau Snopes) sau în seama algoritmilor care analizează conținutul și avertizează dacă sursa are o credibilitate scăzută, astfel de soluții sunt probabil incomplete. În măsura în care astfel de etichete opresc oamenii de la a citi conținutul, pot fi de ajutor; însă, odată ce informația a fost procesată, este improbabil să fie de ajutor o strategie bazată pe evaluarea sursei, mai ales pe măsură ce timpul trece.

În loc de asta, ne concentrăm asupra efectelor fake news și asupra corectării informațiilor false răspândite de articolele din categoria fake news. O problemă este distribuirea corectărilor — nu se poate presupune că oamenii le văd și le procesează. Pe Twitter, cel puțin, distribuirea corecturilor pare mai degrabă descurajantă. Shin și colegii săi au analizat peste 400 000 de tweet-uri referitoare la 57 de zvonuri politice și au ajuns la următoarea concluzie: „În general au existat puține tweet-uri care să fi negat vreun zvon din baza noastră de date, indiferent dacă era adevărat sau fals“ (Shin, Jian, Driscoll și Bar, 2017, p. 1222). Autorii au descoperit că sub 4% dintre tweet-uri negau zvonuri, cele care le susțineau fiind mult mai multe. Și, în măsura în care rezultatele verificării faptelor sunt distribuite, este probabil să fie distribuite cele care confirmă opiniile persoanelor. O examinare a distribuțiilor pe Twitter din 2012 a arătat că oamenii au o probabilitate mai mare de „a da share“ informațiilor provenite din verificarea faptelor (e.g. de la rubrica Fact Checker din Washington Post) care corespund afilierii lor politice. Adică democrații au distribuit informațiile provenite din verificarea faptelor care îl favorizau pe Obama, în timp ce informațiile verificate care îl avantajau pe Romney aveau o probabilitate mai mare de a primi share-uri din partea

republicanilor (Shin și Thorson, 2017). În sfârșit, observăm că, și în cazul în care oamenii văd corectările, s-ar putea ca acestea să nu aibă efect.

Retractările din ziare, de exemplu, nu întotdeauna au urmări; unii americani au continuat să aibă convingeri eronate în legătură cu existența armelor de distrugere în masă în Irak, chiar dacă își aminteau că au citit retractări ale acestor afirmații (Lewandowsky, Stritzke, Oberauer și Morales, 2005).

O întrebare este dacă infirmarea fake news este la fel de dificilă ca infirmarea concepțiilor eronate (e.g. despre legătura dintre vaccinuri și autism). Se știe foarte bine cât de dificil de corectat sunt concepțiile eronate ca legătura dintre vaccinuri și autism (Lewandowsky, Ecker, Seifert, Schwarz și Cook, 2012), corectările adesea pierzându-și eficiența în timp și uneori având chiar un efect de recul, amplificând convingerea (Nyhan și Reifler, 2010; cf. Wood și Porter, 2019). Rămâne deschisă întrebarea dacă articolele din categoria fake news sunt la fel de greu de corectat. Într-un studiu experimental, de exemplu, cercetătorii au examinat convingeri prin afirmații ca „Anthony Scaramucci este investigat de senatul Rusiei” și „Președintele Obama și-a falsificat certificatul de naștere”, afirmații care mai apoi au fost corectate (Porter, Wood și Kirby, 2018). Corectările au reușit destul de bine — deși nu e clar cât crezuseră participanții inițial că articolele erau adevărate (participanții care nu au primit corectările au fost descriși ca „având încredere” în titlurile inițiale). Mai mult, efectele pe termen lung nu au fost examinate și este foarte posibil ca, în timp, corectările să-și piardă eficiența (Butler, Fazio și Marsh, 2011).

Noi considerăm că o cale prudentă este urmarea ideilor generate de literatura amplă și să presupunem că demascarea nu este atât de simplă cum poate părea inițial. Lewandowsky și colegii săi (2012) au subliniat mai multe principii de urmărit în demascare; le trecem în revistă înainte de a oferi câteva recomandări bazate pe propriile noastre cercetări. Cel mai clar sfat este ca, în cadrul corectării, să se evite repetarea afirmației pe care o face articolul din categoria fake news (Lewandowsky et al., 2012). Adică, fiecare mesaj care spune „Pizzagate nu s-a întâmplat” crește fluența cuvântului „Pizzagate”, ceea ce e problematic, dat fiind că amplificarea fluenței unui concept poate crește probabilitatea ca întreaga afirmație să fie etichetată ca adevărată.

Alte recomandări includ utilizarea unui limbaj simplu, ușor de citit, și

adăugarea de fotografii. Fotografiile sunt interesante, oamenii prezentând o tendință de a considera afirmațiile însoțite de fotografii ca fiind mai probabil adevărate decât afirmațiile prezentate fără (Newman, Garry, Bernstein, Kantner și Lindsay, 2012). Adică prezentarea unei imagini cu o țestoasă alături de afirmația „țestoasele sunt surde“ amplifică credința în veridicitatea afirmației în comparație cu condiția în care afirmația „țestoasele sunt surde“ este citită fără a fi însoțită de fotografie (poate pentru că fotografiile îi ajută pe cititori să genereze dovezi care susțin afirmația; Newman et al., 2015). Este interesant de observat că Pennycook și colegii săi au descoperit un efect de adevăr iluzoriu în cazul unor titluri fake news însoțite de fotografii, dat fiind că, în general, așa apar în social media astfel de titluri.

În propria noastră cercetare recentă am identificat o asimetrie în cazul probabilității ca o corectare să inducă o revizuire semnificativă a convingerii. Convingerile eronate au o probabilitate mai mare de a fi corectate atunci când corectarea implică inculcarea unei noi convingeri pozitive, în comparație cu condiția invalidării unei credințe (vezi și Lewandowsky et al., 2012). Specific, oamenii se pricep mai bine să corecteze credințe eronate dacă inițial credeau o negație, ca: „Un poem scris pentru o mireasă nu se numește epitalam“, care apoi a fost corectată prin afirmația „Un poem scris pentru mirească se numește epitalam“. Prin opoziție, credințele eronate sunt mai dificil de corectat dacă credința inițială, falsă, este o afirmație (e.g. „Statuia lui David de Michelangelo se află la Veneția“) și dacă corectarea ia forma unei negații (e.g. „Statuia lui David de Michelangelo nu se află la Veneția“). Această strategie generală de schimbare a credințelor prin oferirea unor noi credințe (afirmații) în loc de anularea credințelor anterioare (negații) poate reprezenta o strategie deosebit de promițătoare pentru combaterea fake news și a dezinformării. De fapt, Horne și colegii săi (2015) au reușit să schimbe unele dintre convingerile și intențiile eronate în privința vaccinării prin faptul de a le oferi indivizilor informații despre pericolele bolilor ca pojarul și oreionul (i.e. inculcând o nouă convingere afirmativă despre riscul de boală), și nu doar prin afirmația că nu există nicio legătură între vaccinuri și autism (i.e. negarea convingerii).

Concluzii și direcții viitoare

Oamenii sunt capabili să asimileze un volum impresionant de informații și pot să rețină multe dintre acestea pe durate foarte mari. În acest capitol am descris modul în care acest sistem are proprietăți care facilitează achiziția de noi informații și permite inferențe și transferuri asupra unor situații noi. De exemplu, sistemul este flexibil și în privința inputurilor (zgomotul de fond), și a outputurilor (depășirea informațiilor aflate pentru generarea de idei noi). Sistemul este eficient din punct de vedere cognitiv, cu scurtături care reflectă starea lumii (e.g. tendința de a presupune că celălalt spune adevărul, bazarea pe fluență ca euristică a evaluării adevărului). În sfârșit, reprezentările tind să fie decontextualizate (fără sursă, actualizate fără senzația de a fi re trăite) într-un mod care facilitează utilizarea acestora în diferite contexte și aplicații.

Dar, după cum am descris aici, proprietățile care oferă avantaje sistemului facilitează, de asemenea, bazarea pe informații false din fake news. Când dau de o sursă fake news, se poate ca oamenii să recunoască sau nu această sursă ca având o credibilitate scăzută. Odată ce încep lectura, oamenii au adesea tendința de a crede că noile informații sunt adevărate; de fapt, s-ar putea să fie necesar să o creadă pentru a o putea înțelege (și să facă un al doilea pas de a nu crede). Cititorii de fake news se bazează pe un volum mare de cunoștințe anterioare când procesează articolul, generând probabil inferențe care depășesc textul afișat. Și se poate să întâlnească în repetate rânduri aceleași afirmații, ceea ce le sporește fluența, sporind, astfel, probabilitatea de a fi crezute ca adevărate.

Aceste chestiuni sunt deosebit de problematice în epoca noastră, pe care unii o numesc „era post-adevărului“ (Lewandowsky, Ecker și Cook, 2017) din cauza volumului de informații false răspândite rapid și la scară largă prin platformele de social media. Aceste platforme pot susține crearea de clici sau de „camere de ecou“ alcătuite din oameni care au aceleași convingeri (în general, eronate) și care caută și distribuie informații care le susțin opiniile. Astfel de camere de ecou au fost identificate pe Twitter (Barberá, Jost, Nagler, Tucker și Bonneau, 2015; Williams, McMurray, Kurz și Lambert, 2015), pe Facebook (Del Vicario et al., 2016) și pe diverse bloguri (Suhay,

Blackwell, Roche și Burggeman, 2015). Se poate ca membrii acestora să nu știe că primesc informații din surse cu credibilitate scăzută; au tendința de a crede aceste informații care vin din partea altor membri ai comunității. Și dacă aceleași afirmații și informații sunt repetate fără a fi corectate, credința că sunt adevărate probabil se va amplifica și va avea potențialul de a duce la noi informații false, dat fiind că oamenii fac inferențe care depășesc ceea ce tocmai s-a spus, generând astfel și mai multă dezinformare.

Probabil este imposibil să oprim răspândirea articolelor fake news și credințele pe care le susțin; în loc de asta, suntem obligați să dezvoltăm strategii viabile de corectare a dezinformării. Astfel de eforturi nu trebuie să se concentreze asupra sursei dezinformării inițiale, ci recomandarea noastră este să ne concentrăm asupra modului de transmitere a mesajului: conținutul trebuie să fie simplu și însoțit de imagini (pentru a amplifica fluenta), trebuie să evite să facă judecăți de valoare și trebuie să afirme informația corectă în loc de a nega mitul (și, în general, miturile nu ar trebui formulate). Mai mult, trebuie luată în considerare propagarea mesajelor de retractare/de contrazicere, dat fiind că oamenii sunt mult mai interesați de afirmațiile inițiale decât de corectări.

În acest capitol am susținut că, pentru combaterea erorilor din sistemul de cunoștințe (convingeri eronate, iluzia cunoașterii, concepții greșite), s-ar putea să fie nevoie de strategii diferite decât cele recomandate pentru prevenirea amintirii eronate a evenimentelor specifice. Se poate însă să existe abordări viabile ale corectării care depășesc centrarea pe sursă și care se folosesc de memoria episodică. De exemplu, un eveniment deosebit de memorabil (e.g. o conferință de presă încărcată afectiv la care persoana participă direct) în care o afirmație falsă este corectată (sau retractată) poate avea o probabilitate mai mare de a provoca o schimbare durabilă a unei convingeri decât un eveniment mai puțin memorabil care implică o corectare. În acest mod, cunoștințele și memoria episodică pot interacționa în anumite contexte pentru a spori probabilitatea ca o corectare să se mențină în timp. Este nevoie de mai multe cercetări pentru verificarea durabilității corectării în cazul acestui tip de interacțiune.

Note

1.Desigur, se poate ca o convingere eronată să fie provocată de o amintire falsă, după cum demonstrează Polage (2012); dar noi considerăm că această situație este mult mai puțin probabilă decât iluziile de cunoaștere.

2.Deși controversate, cercetări mai recente susțin o poziție diferită, în speță că validarea poate uneori să susțină înțelegerea (Richter, 2015; Richter, Schroeder și Wöhrmann, 2009). Din această perspectivă, validarea informațiilor noi tinde să fie un proces relativ automat și fără efort care are loc fără înțelegere, și nu un pas distinct, care presupune efort și care apare după ce informația a fost înțeleasă.

Referințe

Barberá, P., Jost, J.T., Nagler, J., Tucker, J.A., Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26 (10), 1531–1542.

Bauer, P.J. și San Souci, P. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental Child Psychology*, 107, 452–465.

Begg, I.M., Anas, A. și Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, 121 (4), 446.

Begg, I.M., și Armour, V. (1991). Repetition and the ring of truth: Biasing comments. *Canadian Journal of Behavioural Science*, 23 (2), 195.

Bottoms, H.C., Eslick, A.N. și Marsh, E.J. (2010). Memory and the Moses illusion: Failures to detect contradictions with stored knowledge yield negative memorial consequences. *Memory*, 18 (6), 670–678.

Brashier, N.M., Eliseev, E.D. și Marsh, E.J. (2020). An initial accuracy focus prevents illusory truth. *Cognition*, 194, 104054.

Brewer, W.F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition*, 5, 673–678.

Bronstein, M.V., Pennycook, G., Bear, A., Rand, D.G. și Cannon, T.D. (2019). Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *Journal of Applied Research in Memory and Cognition*, 8 (1), 108–117.

Brown, A.S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, 109, 204–223.

Butler, A.C., Fazio, L.K. și Marsh, E.J. (2011). The hypercorrection effect persists over a week, but high-confidence errors return. *Psychonomic Bulletin & Review*, 18, 1238–1244.

Cantor, A.D., Eslick, A.N., Marsh, E.J., Bjork, R.A. și Bjork, E.L. (2015). Multiple-choice tests stabilize access to marginal knowledge. *Memory & Cognition*, 43, 193–205.

Cantor, A.D. și Marsh, E.J. (2017). Expertise effects in the Moses illusion: Detecting contradictions with stored knowledge. *Memory*, 25, 220–230.

Conway, M.A., Gardiner, J.M., Perfect, T.J., Anderson, S.J. și Cohen, G.M. (1997). Changes in memory awareness during learning: The acquisition of knowledge by psychology undergraduates. *Journal of Experimental Psychology: General*, 126, 393–413.

Dechêne, A., Stahl, C., Hansen, J. și Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, 14, 238–257.

Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G. și Quattrociocchi, W. (2016). Echo chambers: Emotional contagion and group polarization on facebook. *Scientific Reports*, 6, 37825.

DePaulo, B.M., Kashy, D.A., Kirkendol, S.E., Wyer, M.M. și Epstein, J.A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, 70 (5), 979–995.

Erickson, T.D. și Mattson, M.E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning & Verbal Behavior*, 20, 540–551.

Eslick, A.N., Fazio, L.K. și Marsh, E.J. (2011). Ironic effects of drawing attention to story errors. *Memory*, 19 (2), 184–191.

Fazio, L.K., Brashier, N.M., Payne, B.K. și Marsh, E.J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General*, 144 (5), 993–1002.

Fazio, L.K., Rand, D.G. și Pennycook, G. (2019). Repetition increases perceived truth equally for plausible and implausible statements. *Psychon Bull Rev* 26, 1705-1710.

Ferreira, F., Bailey, K.G. și Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11, 11–15.

Fragale, A.R. și Heath, C. (2004). Evolving informational credentials: The (mis)attribution of believable facts to credible sources. *Personality and Social Psychology Bulletin*, 30 (2), 225–236.

Gilbert, D.T. (1991). How mental systems believe. *American Psychologist*, 46, 107–119.

Gilbert, D.T., Krull, D.S. și Malone, P.S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology*, 59, 601–613.

Gilbert, D.T., Tafarodi, R.W., și Malone, P.S. (1993). You can't not believe

everything you read. *Journal of Personality and Social Psychology*, 65, 221–233.

Guillory, J.J. și Geraci, L. (2013). Correcting erroneous inferences in memory: The role of source credibility. *Journal of Applied Research in Memory and Cognition*, 2 (4), 201–209.

Hasher, L., Goldstein, D. și Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16, 107–112.

Henkel, L.A. și Mattson, M.E. (2011). Reading is believing: The truth effect and source credibility. *Consciousness and Cognition*, 20 (4), 1705–1721.

Horne, Z., Powell, D., Hummel, J.E. și Holyoak, K.J. (2015). Countering antivaccination attitudes. *Proceedings of the National Academy of Sciences*, 112 (33), 10321–10324.

Isberner, M.B. și Richter, T. (2014). Does validation during language comprehension depend on an evaluative mindset? *Discourse Processes*, 51 (1–2), 7–25.

Johnson, M.K., Hashtroudi, S. și Lindsay, D.S. (1993). Source monitoring. *Psychological Bulletin*, 114, 3–28.

Kornell, N. și Bjork, R.A. (2008). Learning concepts and categories: Is spacing the „enemy of induction“? *Psychological Science*, 19 (6), 585–592.

Lewandowsky, S., Ecker, U.K. și Cook, J. (2017). Beyond misinformation: Understanding and coping with the „post-truth“ era. *Journal of Applied Research in Memory and Cognition*, 6 (4), 353–369.

Lewandowsky, S., Ecker, U.K., Seifert, C.M., Schwarz, N. și Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13, 106–131.

Lewandowsky, S., Stritzke, W.G., Oberauer, K. și Morales, M. (2005). Memory for fact, fiction, and misinformation: The Iraq War 2003.

Psychological Science, 16, 190–195.

Lindsay, D.S. și Johnson, M.K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, 17, 349–358.

Mandelbaum, E. (2014). Thinking is believing. *Inquiry: An Interdisciplinary Journal of Philosophy*, 57 (1), 55–96.

Marsh, E.J., Cantor, A.D. și Brashier, N.M. (2016). Believing that humans swallow spiders in their sleep: False beliefs as side effects of the processes that support accurate knowledge. *Psychology of Learning and Motivation*, 64, 93–132.

Marsh, E.J., Meade, M.L. și Roediger III, H.L. (2003). Learning facts from fiction. *Journal of Memory and Language*, 49 (4), 519–536.

Marsh, E.J. și Yang, B.W. (2017). A call to think broadly about information literacy. *Journal of Applied Research in Memory and Cognition*, 6, 401–404.

McGlone, M.S. și Tofighbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, 11, 424–428.

Mills, J. și Jellison, J.M. (1967). Effect on opinion change of how desirable the communication is to the audience the communicator addressed. *Journal of Personality and Social Psychology*, 6 (1), 98.

Newman, E.J., Garry, M., Bernstein, D.M., Kantner, J. și Lindsay, D.S. (2012). Nonprobative photographs (or words) inf late truthiness. *Psychonomic Bulletin & Review*, 19, 969–974.

Newman, E.J., Garry, M., Unkelbach, C., Bernstein, D.M., Lindsay, D.S. și Nash, R.A. (2015). Truthiness and falsiness of trivia claims depend on judgmental contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41, 1337–1348.

Nyhan, B. și Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32, 303–330.

Pennycook, G., Cannon, T.D. și Rand, D.G. (2018). Prior exposure increases

perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147 (12), 1865–1880.

Pennycook, G. și Rand, D.G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39–50.

Polage, D.C. (2012). Making up history: False memories of fake news stories. *Europe's Journal of Psychology*, 8, 245–250.

Porter, E., Wood, T.J. și Kirby, D. (2018). Sex trafficking, Russian infiltration, birth certificates, and Pedophilia: A survey experiment correcting fake news. *Journal of Experimental Political Science*, 5, 159–164.

Priester, J.R. și Petty, R.E. (1995). Source attributions and persuasion: Perceived honesty as a determinant of message scrutiny. *Personality and Social Psychology Bulletin*, 21 (6), 637–654.

Putnam, A.L. și Phelps, R.J. (2017). The citation effect: In-text citations moderately increase belief in trivia claims. *Acta Psychologica*, 179, 114–123.

Rapp, D.N. și Donovan, A.M. (2017). Routine processes of cognition result in routine influences of inaccurate content. *Journal of Applied Research in Memory and Cognition*, 6 (4), 409–413.

Rapp, D.N., Hinze, S.R., Kohlhepp, K. și Ryskin, R.A. (2014). Reducing reliance on inaccurate information. *Memory & Cognition*, 42, 11–26.

Rapp, D.N. și Salovich, N.A. (2018). Can't we just disregard fake news? The consequences of exposure to inaccurate information. *Policy Insights from the Behavioral and Brain Sciences*, 5 (2), 232–239.

Reber, R. și Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, 8, 338–342.

Richter, T. (2015). Validation and comprehension of text information: Two sides of the same coin. *Discourse Processes*, 52 (5–6), 337–355.

Richter, T., Schroeder, S. și Wöhrmann, B. (2009). You don't have to believe

everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology*, 96 (3), 538.

Shin, J., Jian, L., Driscoll, K. și Bar, F. (2017). Political rumoring on Twitter during the 2012 US presidential election: Rumor diffusion and correction. *New Media & Society*, 19, 1214–1235.

Shin, J., Jian, L., Driscoll, K. și Bar, F. (2018). The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior*, 83, 278–287.

Shin, J. și Thorson, K. (2017). Partisan selective sharing: The biased diffusion of fact-checking messages on social media. *Journal of Communication*, 67, 233–255.

Stanley, M.L., Yang, B.W. și Marsh, E.J. (2019). When the unlikely becomes likely: Qualifying language does not influence later truth judgments. *Journal of Applied Research in Memory and Cognition*, 8 (1), 118–129.

Street, C.N. și Kingstone, A. (2017). Aligning Spinoza with Descartes: An informed Cartesian account of the truth bias. *British Journal of Psychology*, 108 (3), 453–466.

Street, C.N. și Richardson, D.C. (2015). Descartes versus Spinoza: Truth, uncertainty, and bias. *Social Cognition*, 33 (3), 227–239.

Strickland, B. și Keil, F. (2011). Event completion: Event based inferences distort memory in a matter of seconds. *Cognition*, 121, 409–415.

Suhay, E., Blackwell, A., Roche, C., și Bruggeman, L. (2015). Forging bonds and burning bridges: Polarization and incivility in blog discussions about Occupy Wall Street. *American Politics Research*, 43 (4), 643–679.

Swami, V., Voracek, M., Stieger, S., Tran, U.S. și Furnham, A. (2014). Analytic thinking reduces belief in conspiracy theories. *Cognition*, 133, 572–585.

Swire, B. și Ecker, U.K.H. (2018). Misinformation and its correction: Cognitive mechanisms and recommendations for mass communication. *Misinformation and Mass Audiences*, 195–211.

Tulving, E. (1972). Episodic and semantic memory. În E. Tulving și W. Donaldson (coord.), *Organization of memory* (pp. 381–403). New York, NY: Academic Press.

Underwood, J. și Pezdek, K. (1998). Memory suggestibility as an example of the sleeper effect. *Psychonomic Bulletin & Review*, 5, 449–453.

Watkins, M.J. și Kerkar, S.P. (1985). Recall of a twice-presented item without recall of either presentation: Generic memory for events. *Journal of Memory and Language*, 24, 666–678.

Williams, H.T., McMurray, J.R., Kurz, T. și Lambert, F.H. (2015). Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Environmental Change*, 32, 126–138.

Wood, T. și Porter, E. (2019). The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior*, 41, 135–163.

Capitolul 9. Inoculare psihologică împotriva fake news

Sander van der Linden și Jon Roozenbeek

Mulțumiri

Dorim să mulțumim Cambridge Social-Decision-Making Lab, Ruurd Oosterwoud și DROG, Design Studio Gusmanson și Economic and Social Research Council (ESRC) și Universității din Cambridge pentru susținere și pentru finanțarea cercetării. Le mulțumim, de asemenea, editorilor și evaluatorilor pentru feedbackul oferit pentru o versiune anterioară a acestui capitol.

Ce înseamnă fake news?

Fake news sunt peste tot. După ce a luat avânt în timpul campaniei electorale din 2016 din SUA, această sintagmă a devenit omniprezentă în media populare. Președintele SUA Donald Trump o folosește pentru a ocări jurnaliștii și instituțiile de presă pentru știri pe care le consideră tendențioase; cercetătorii creează algoritmi de detectare a articolelor false și eronate și le consemnează răspândirea; platforma Facebook este constant

obligată să explice cum are de gând să împiedice articolele fake news să ajungă virale; și guvernele iau măsuri pentru a combate articolele fake news care circulă pe internet (Bremner, 2018; Shao et al., 2018; Wakabayashi, 2017).

Astfel, „fake news“ a ajuns rapid o sintagmă generică, fără o definiție de lucru acceptată (Tandoc Jr., Lim și Ling, 2018). Deși unii au încercat să explice „știința fake news“ (e.g. Lazer et al., 2018), dacă oameni și instituții, de la BBC la președintele Donald Trump, o folosesc în sensuri complet diferite în contexte diferite, e dificil de explicat la ce ne referim când vorbim despre fake news.

Este, de asemenea, destul de clar că sintagma nu e foarte potrivită pentru descrierea întregii probleme. Poate o definiție rezonabilă a „fake news“ ar putea fi „informații inventate care imită conținutul publicat de instituțiile media în ce privește forma, dar nu și procesul organizatoric sau intenția“ (Lazer et al., 2018, p. 1). Snopes este unul dintre website-urile care urmăresc astfel de articole. Exemple nu sunt greu de găsit: titluri ca „Cetățenii australieni vor fi vaccinați prin substanțe răspândite în aer“, „Melania Trump interzice personalului de la Casa Albă să se vaccineze antigripal“ și „Medic musulman refuză să trateze o creștină la bordul avionului“ nu sunt decât la o căutare distanță pe Google (Adl-Tabatabai, 2016; Baxter, 2018; Patriot United, 2018).

Dar articolele de știri nu trebuie să fie complet false pentru a induce în eroare. E simplu ca lucrurile pe care le spun oamenii să fie scoase din context pentru a da impresia că au spus ceva ce de fapt n-au spus și e simplu ca unui montaj video sau unei imagini să i se adauge un context care schimbă datele. De exemplu, știrea din Figura 9.1, comentariu postat de pagina de Facebook „News World“ în data de 20 martie 2018, care pretinde că imigranți musulmani din Franța asaltează o biserică catolică în timpul slujbei. Înregistrarea a fost văzută de aproximativ 1,2 milioane de ori într-o zi de când a fost postată. Politicienii, printre care Marine le Pen, lidera Frontului Național, și-au exprimat indignarea pe Twitter, scriind că biserica a fost „pângărită“ (Le Pen, 2018).



News World

...

Yesterday at 6:35am · 🌐



1.2M Views

Figura 9.1. Postare pe Facebook de către News World (20 martie 2018)

Saint-Denis (France): Imigranți musulmani asaltează o biserică catolică în timpul slujbei. Poliția încearcă să îi oprească

Dar verificarea faptelor a arătat rapid că există mai multe probleme în privința acestor afirmații. Nu exista nicio dovadă a religiei protestatarilor sau a perioadei în care au ajuns în Franța. Mai mult, biserica nu a fost „asaltată”, cel puțin după spusele enoriașilor înșiși. Persoanele din înregistrare protestau împotriva unui proiect de lege care ar fi făcut mai dificilă obținerea azilului politic în țară (Snopes, 2018). Demonstrația a rămas nonviolentă.

Problema cu postarea pe Facebook din Figura 9.1 nu este că informația e complet falsă sau că evenimentele prezentate în înregistrare nu s-au întâmplat. Contextul fals oferit în postare este adevărata problemă. Pe scurt, dintr-o perspectivă psihologică, intenția contează și eronat și contrafăcut nu înseamnă chiar același lucru. Merită, așadar, să ne gândim la tipurile diferite de fake news ca aflându-se pe un spectru. La un capăt al spectrului avem informațiile false (misinformation), adică informație pur și simplu falsă sau greșită (și poate fi cauzată inclusiv de greșeli omenești). Apoi avem dezinformare (disinformation), care include informații false și intenția de a înșela audiența. În comparație cu simpla greșală omenească, intenția are conotații psihologice importante (van der Linden, 2017; van der Linden, Roozenbeek, Oosterwoud, Compton și Lewandowsky, 2018). Propaganda este astfel definită ca dezinformare pe baza unei agende politice explicite sau implicite (van der Linden, 2017). Pentru a nu complica lucrurile, în acest capitol vom folosi termenul dezinformare în loc de fake news, astfel încât să fie clar că nu ne referim doar la articole fake news, ci la manipularea media în general.

Problema dezinformării

Deși este evident că nu e un fenomen nou, dezinformarea a devenit o problemă mult mai gravă odată cu apariția internetului. Posibilitatea comunicării instantanee și anonime face din internet un vehicul ideal pentru înșelare. Astfel, poate nu e surprinzător faptul că dezinformarea e folosită adesea de diferiți actori, inclusiv de unele guverne, pentru a influența opinia publică. Platformele de social media sunt un teren deosebit de fertil. Pentru a oferi un exemplu, în jur de 47 de milioane de conturi Twitter (aproximativ 15%) sunt boți (conturi automatizate ruse) (Varol, Ferrara, Davis, Menczer și Flammini, 2017) și mulți dintre acești boți sunt folosiți pentru a răspândi intenționat dezinformare politică, în special în timpul campaniilor electorale (Ferrara, 2017).

Printre exemplele recente de dezinformare la scară largă se numără teoriile conspirației referitoare la COVID-19, MacronLeaks din timpul alegerilor prezidențiale din 2017 din Franța, controversa Pizzagate din timpul alegerilor din 2016 din SUA, diferitele explicații „alternative” ale prăbușirii avionului MH17 al Malaysia Airlines din iulie 2014 și zvonurile care circulă în Suedia despre cooperarea țării cu NATO (Kragh și Åsberg, 2017). Această abundență de dezinformare online își spune cuvântul. De exemplu, să ne gândim că majoritatea americanilor recunosc că din cauza fake news sunt derutați în privința unor fapte banale (Barthel, Mitchell și Holcomb, 2016) și 83% dintre europeni cred că fake news reprezintă o amenințare la adresa democrației (Eurobarometru, 2018). Mai mult, un studiu recent realizat în Marea Britanie de YouGov a arătat că doar 4% dintre participanți au putut să facă diferența între fake news și știri reale (Channel 4, 2017). Sub unele aspecte, aceste lucruri nu sunt surprinzătoare: oamenii sunt bombardați cu informații tot timpul cât se uită prin feed-urile de știri.

Dat fiind că orice democrație se bazează pe o populație bine informată, răspândirea dezinformării se arată a fi o amenințare reală la adresa procesului democratic (Lewandowsky, Ecker și Cook, 2017; van der Linden et al., 2017). În multe țări, și tabloul politic devine din ce în ce mai polarizat. Acest nivel de polarizare este susținut de apariția bulelor omogene din punct de vedere ideologic, oamenii fiind expuși la articole compatibile perspectivei lor ideologice (Del Vicario et al., 2016). Dacă un număr mare de oameni nu

știi ce să creadă și în cine să aibă încredere, asta poate avea consecințe serioase pentru luarea de decizii bazate pe dovezi într-un spectru larg de chestiuni, de la încălzirea globală și vaccinuri la relații internaționale (Lewandowsky et al., 2017; Poland și Spier, 2010; van der Linden, Leiserowitz, Rosenthal și Maibach, 2017). În unele cazuri, răspândirea virală de articole din categoria fake news a dus la răniri și chiar la moarte (BBC, 2018a).

Știm, de asemenea, că nu e ușor de anulat efectul campaniilor eficiente de dezinformare. Studii asupra influenței constante a dezinformării arată consecvent că, odată asimilate, convingerile sunt foarte dificil de corectat, chiar dacă oamenii recunosc că perspectivele lor se bazează pe informații eronate (Lewandowsky, Ecker, Seifert, Schwarz și Cook, 2012). Astfel, deși dezinformarea este probabil la fel de veche ca înseși zvonurile, crearea și diseminarea dezinformării este acum mai simplă decât oricând și din ce în ce mai mulți oameni sunt expuși la conținut fals pur și simplu prin consumul zilnic de știri. Așadar, guvernele, autoritățile publice și companiile media au propus cu toții diferite soluții potențiale pentru combarea problemei în continuă creștere a dezinformării.

Soluții potențiale

Soluțiile propuse pentru rezolvarea problemei dezinformării pot fi împărțite în patru mari categorii: (1) algoritmice, (2) corective, (3) legislative și (4) psihologice. În prima categorie, Google și Facebook descoperă moduri de ajustare a algoritmilor, astfel încât să descurajeze site-urile de știri false sau îndoielnice și să împiedice informațiile false să apară în feed-urile utilizatorilor așa cum apar articolele site-urilor „de încredere” (Calfas, 2017; Elgin și Wang, 2018). Cu toate acestea, este evident că algoritmii nu funcționează perfect și nu reușesc să identifice conținuturile înșelătoare și diferite încercări, cum ar fi cele făcute de Facebook, au avut adesea efecte secundare (Wakefield, 2017). A doua categorie se referă la corectarea post-

hoc a știrilor false prin instrumente de verificare a faptelor. Sunt multe inițiative de verificare a faptelor, și unele (cum ar fi PolitiFact și Snopes) sunt consacrate. Însă, deși aceste inițiative sunt lăudabile, dovezile în ce privește eficacitatea lor sunt amestecate (Nyhan, Porter, Reifler și Wood, 2019). Mai mult, este imposibilă demascarea fiecărui articol fals sau eronat, iar audiența rapoartelor de verificare rămâne limitată (Kurzleben, 2016). Cercetări recente indică și faptul că știrile false se răspândesc mai ușor pe social media decât orice alt tip de știre, chiar și în condițiile în care se controlează știrile distribuite de boții ruși (Vosoughi, Roy și Aral, 2018).

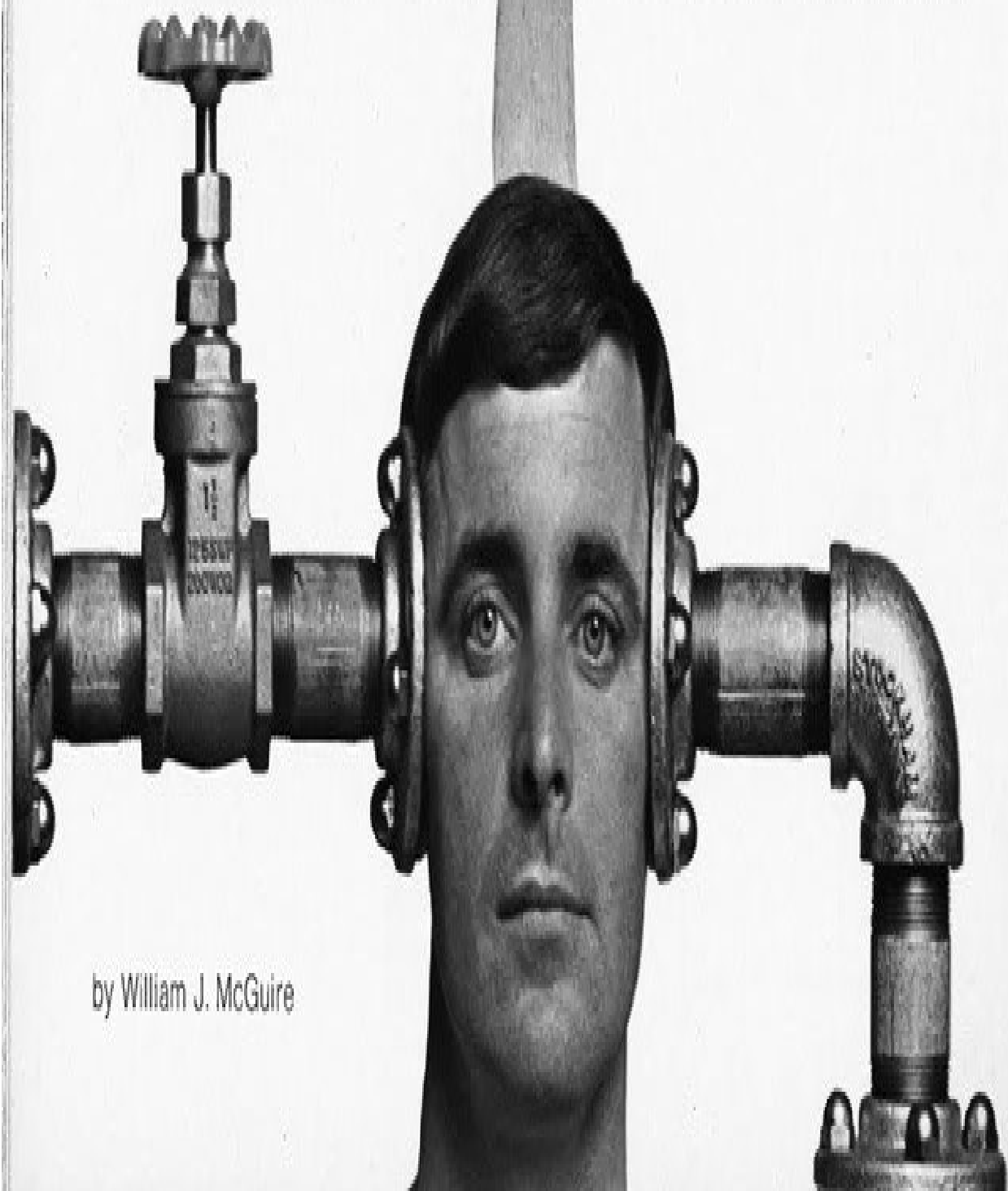
O abordare mai radicală presupune introducerea unei noi reglementări și legislații pentru combaterea dezinformării. Un exemplu proeminent este „Legea împotriva știrilor false“ din Franța, conform căreia în timpul alegerilor se impun restricții mai dure în privința conținutului pe care instituțiile media le pot publica online (Bremner, 2018). Inițiative similare au fost propuse în Marea Britanie (e.g. „comisia pentru reglementarea știrilor“ (BBC News, 2018b). Însă împuternicirea oricărei organizații, guvernamentale sau nu, de a decide care informații sunt „reale“ și care sunt „false“ poate avea rapid efecte negative. De exemplu, un grup de lucru finanțat de UE, numit „EUvsDisinfo“ a fost aspru criticat pentru a fi etichetat mai multe site-uri olandeze nu foarte populare și un ziar local ca „difuzând dezinformare rusească“ (Pieters, 2018). Parlamentari olandezi și-au exprimat îngrijorarea în privința faptului că EUvsDisinfo încălca dreptul la liberă exprimare și au votat desființarea grupului.

Idei oferite de psihologie: vaccinare împotriva dezinformării

Așadar, atenția se îndreaptă acum mai mult spre rolul psihologiei, educației și științelor comportamentale pentru combaterea fake news prin susținerea populației la nivel individual (Comisia Europeană, 2018). Desigur, este rezonabilă ideea de bază că dezinformarea va fi mai puțin eficientă dacă vom

avea consumatori de știri care gândesc critic și sunt bine informați. Însă problema majorității abordărilor tradiționale de evaluare critică a presei stă în faptul că nu e nicidecum posibilă corectarea tuturor știrilor false. O altă problemă esențială este că dezvoltarea unor tehnici mai bune de demascare este improbabil să fie o soluție suficientă împotriva nenumăratelor fake news. De fapt, chiar dacă sunt publicate corectări, paguba va fi fost deja făcută: odată ce oamenii au asimilat o convingere eronată, este improbabil să-și actualizeze perspectiva. Într-adevăr, cercetarea asupra „efectului influenței continue“ sugerează că corectările sunt adesea ineficiente, oamenii continuând să se bazeze pe informații false (și demascate), în special în cazul convingerilor motivate politic (Lewandowsky et al., 2012). Ceea ce aduce în discuție întrebarea: este posibilă prevenția răspândirii de narațiuni false? Pentru a răspunde la această întrebare, apelăm la ceea ce Eagly și Chaiken (1993, p. 561) au numit „teoria originară a rezistenței la persuasiune“: teoria inoculării. În timpul celui de-al Doilea Război Mondial, Ministerul de Război al SUA avea o subdiviziune de cercetare experimentală în cadrul unei unități numite „Departamentul de informare și educație“. Această divizie, condusă de Carl Hovland, specialist în psihologie socială, avea sarcina specifică de a realiza cercetări asupra campaniilor de persuasiune și propagandă politică (Hovland, Lumsdaine și Sheffield, 1949). Aceste studii au constituit baza a ceea ce mai târziu a ajuns cunoscut ca unul dintre primele grupuri de psihologie socială: Programul Yale de Atitudini și Persuasiune (Yale Attitude and Persuasion Program) (Huddy, Sears și Levy, 2013). Misiunea lor este să descopere „legile fundamentale ale persuasiunii“ folosind metode experimentale controlate. Însă în anii 1960 a apărut o nouă chestiune: potențiala spălare pe creier a soldaților americani capturați în Orientul Îndepărtat. Așadar, psihologul William McGuire a abordat o întrebare diferită: cum putem ajuta oamenii să reziste încercărilor de persuasiune? Eforturile sale au dus în cele din urmă la dezvoltarea „teoriei inoculării“, pe care a descris-o ca „vaccin pentru spălarea pe creier“ (vezi Figura 9.2).

A Vaccine for **BRAINWASH**



by William J. McGuire

Figura 9.2. Un vaccin pentru spălarea pe creier (McGuire, 1970)

Teoria inoculării se bazează pe o analogie cu procedura imunologică (McGuire, 1970; McGuire și Papageorgis, 1961, 1962; Papageorgis și McGuire, 1961). Vaccinurile sunt versiuni atenuate ale patogenilor (e.g. un virus), care, prin introducere în corp, declanșează producerea de anticorpi. Acești anticorpi devin activi odată ce versiunea reală a patogenului ajunge în organism, conferind astfel protecție (imunitate) împotriva infecției ulterioare. Teoria inoculării postulează că același lucru se întâmplă în cazul informației: prezentarea preventivă a unei versiuni atenuate de informație eronată declanșează un proces cognitiv analog cultivării de „anticorpi mentali”, făcând persoana imună la încercări (nedorite) de persuasiune (Compton, 2013; McGuire și Papageorgis, 1961). De-a lungul anilor au fost strânse din ce în ce mai multe dovezi care arată că inocularea funcționează în cazul atitudinilor publicului în privința mai multor domenii, printre care sănătatea (Compton, Jackson și Dimmock, 2016) și politica (Pfau, Park, Holbert și Cho, 2001). Și metaanalizele confirmă faptul că mesajele de inoculare sunt eficiente și conferă rezistență la persuasiune (Banas și Rains, 2010). Procesul de inoculare constă în două componente esențiale, și anume: (1) un avertisment care activează perceperea amenințării de către receptori (baza afectivă) și (2) refutare preventivă (baza cognitivă). Se consideră că avertizarea oamenilor că urmează să fie expuși la conținuturi contraatitudinale provoacă activarea percepției amenințării, care motivează protejarea convingerilor existente. La rândul lor, mesajele de refutare bidimensionale informează și totodată educă, în sensul că modelează procesul de contraargumentare și oferă un conținut specific care poate fi utilizat pentru a rezista la încercările de persuasiune (McGuire, 1970; Compton, 2013).

Este interesant că în privința acestei teorii rămân mai multe întrebări deschise, în special în ce privește aplicarea acesteia la fake news și dezinformare. De exemplu, teoria inoculării a fost aplicată în general așa-numitelor truisme culturale sau convingerilor larg răspândite (e.g. convingerea că spălarea pe dinți de două ori pe zi este o practică sănătoasă, vezi McGuire, 1970). Astfel, o întrebare majoră este: Cum operează în teorie procesul inoculării în cazul în care oamenii au atitudini inițiale divergente în

ce privește o anumită chestiune (după cum se întâmplă adesea în cazul fake news). Dacă audiența nu are deja atitudinea dorită, procesul de inoculare nu este profilactic în sensul tradițional, ci ia mai degrabă un rol „terapeutic” — analog apariției vaccinurilor terapeutice (Compton, 2019). În al doilea rând, dintr-o perspectivă a aplicării, rămâne neclar modul în care procesul inoculării poate fi adus la nivelul întregii populații, dat fiind că e evident nefezabilă refutarea preventivă a fiecărui articol de fake news în parte. În sfârșit, tratamentele prin inoculare se bazează în general pe un proces „pasiv”, prin care receptorii citesc un mesaj persuasiv care îi avertizează și respinge contraargumente potențiale. Cu toate astea, McGuire a teoretizat de la început că un proces activ de inoculare ar putea avea un efect mai puternic în cazul în care oamenii generează propriile argumente pro și contra.¹ Astfel, am realizat trei studii prin care am încercat să oferim primele răspunsuri la aceste întrebări importante și încă nesoluționate.

În laborator: inocularea publicului împotriva dezinformării cu privire la încălzirea globală (studiul 1)

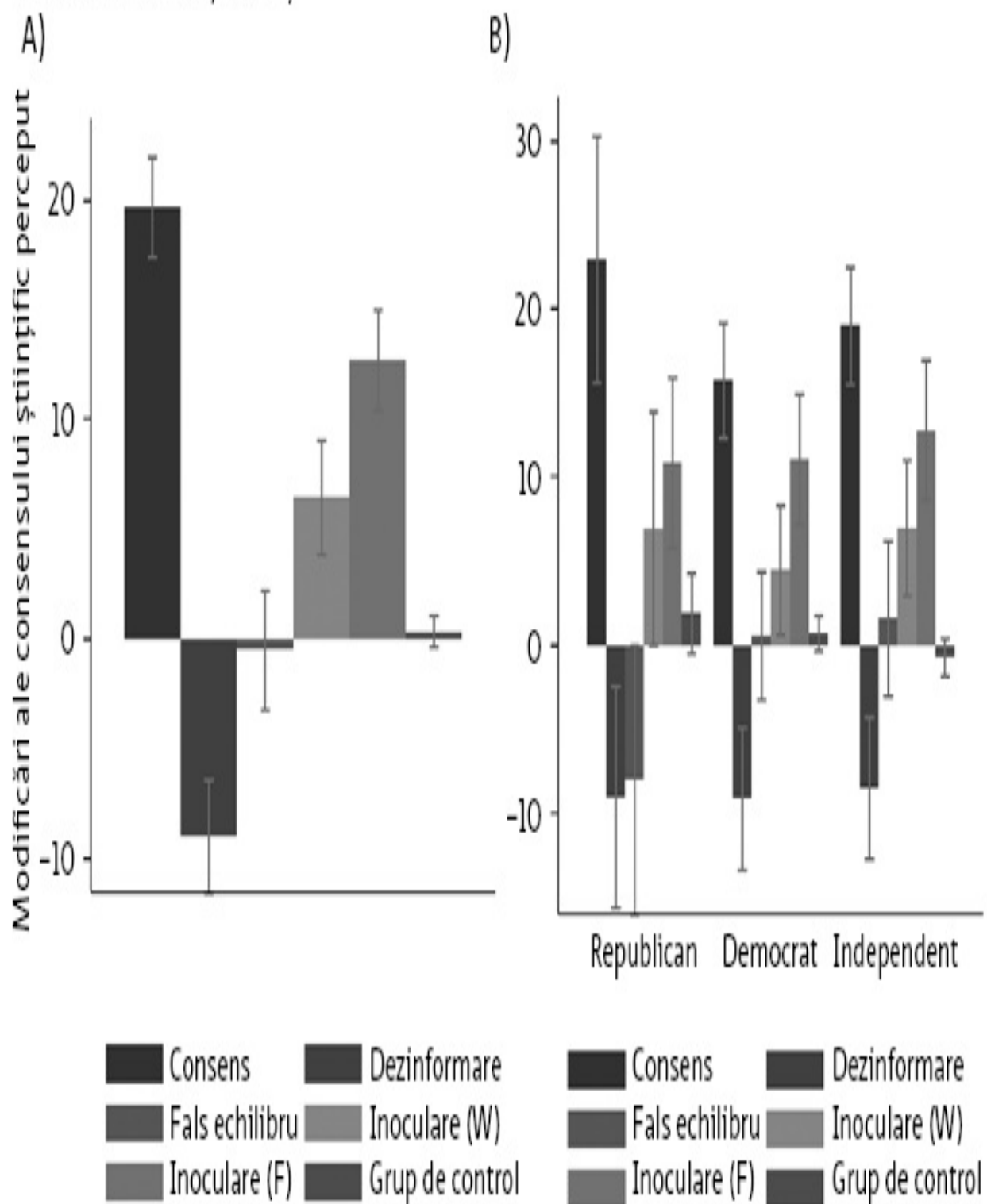
Pentru a răspunde la prima întrebare, am dorit să vedem dacă expunerea publicului la o versiune atenuată a informațiilor false și demascarea preventivă a falsului prin fapte demonstrate științific (un vaccin) poate determina rezistența în fața fake news referitoare la o chestiune foarte polarizată și controversată: încălzirea globală. Am realizat în laborator două mari studii online pentru a testa aceste ipoteze. În primul studiu am folosit un eșantion selectat aleatoriu (N = 1 000) de cetățeni SUA pentru a evalua care sunt „falsurile” cele mai familiare pentru oameni în privința chestiunii respective. Sursa de dezinformare care a avut cea mai mare recunoaștere a fost o petiție online reală (Petiția Oregon), care susține că a strâns peste 31 000 de semnături de la oameni de știință care nu sunt de acord că încălzirea globală este reală (alimentând cel mai popular articol fake news

despre încălzirea globală de pe social media din anul 2016, vezi Readfearn, 2016).

În cadrul celui de-al doilea studiu, ne-am bazat pe un eșantion mare și divers ($N = 2\,167$) de adulți americani recrutați prin Amazon Mechanical Turk. Am folosit un screenshot real al website-ului de petiții în condiția experimentală de „dezinformare” și, ca afirmație „factuală”, simplul fapt că peste 97% dintre specialiștii în climatologie au ajuns la concluzia că încălzirea globală provocată de oameni este reală (Cook et al., 2016). Apoi am alocat aleatoriu participanții într-una din șase condiții și le-am cerut să evalueze consensul științific și înainte, și după (vezi Figura 9.4). Cele șase condiții au fost: (1) fapte concrete (consensul de 97%), (2) dezinformare concretă (petiția), (3) fals echilibru (consens versus petiție), (4) inoculare parțială (doar avertisment inițial), (5) inoculare completă (avertisment inițial plus refutare preventivă) și (6) grup de control „pur”, în care participanții au rezolvat un puzzle neutru de cuvinte.

În Figura 9.3, de la stânga la dreapta (graficul a), în condiția „fapte”, participanții doar au citit informații despre dovezi și și-au schimbat perspectivele (pre-post) asupra încălzirii globale conform consensului științific (97%) în proporție de aproximativ 20% ($d = 1,23$). În condiția „dezinformare”, subiecților li s-a arătat doar informația falsă și și-au schimbat perspectiva în proporție de aproximativ 10% ($d = 0,48$). În condiția „fals echilibru”, participanții au văzut ambele versiuni una lângă alta; aici, prezența „constantă” a dezinformării a neutralizat complet faptele ($d = 0,04$). În ultimele două condiții, de „inoculare”, participanții au fost informați cu privire la fapte fie printr-un scurt avertisment (W (warning)) că s-ar putea ca actori cu motivații politice să influențeze opinia publică prin tactici de inducere în eroare sau au fost folosite mai multe argumente pentru a demasca preventiv și în detaliu dezinformația (e.g. printre semnatarii petiției erau Charles Darwin și Spice Girls). Condițiile de inoculare (etichetate „W” de la warning (avertisment) și „F” de la full (toate) în Figura 9.3) s-au dovedit eficiente, reținând aproximativ o treime ($d = 0,33$), respectiv două treimi ($d = 0,75$) din mesajul concret. Este esențial faptul că același tipar s-a menținut pe întregul spectru politic, precum și în privința atitudinilor anterioare ale participanților cu privire la încălzirea globală (graficul b).

Figura 9.3. Inocularea împotriva dezinformării în privința încălzirii globale (van der Linden et al., 2017)



Pe scurt, acest studiu a oferit un răspuns important și credibil la prima întrebare: chiar și în afara contextului „truismelor” neutre din punct de vedere politic, inocularea poate fi eficientă. De fapt, în loc să provoace efecte secundare, inocularea pare să aibă beneficii retroactive sau „terapeutice”, chiar și în cazul oamenilor care resping perspectiva științifică asupra încălzirii globale din motive politice. Un studiu independent a reprodus conceptual aceste descoperiri, cu rezultate în general mai accentuate (vezi Cook, Lewandowsky și Ecker, 2017).

Pe teren: inoculare activă împotriva fake news (studiul 2)

Cercetarea noastră inițială ne-a lăsat cu două întrebări importante: (1) în loc de citirea pasivă de articole, este posibilă inocularea oamenilor într-o manieră mai degrabă „activă” și „experiențială” care să susțină procesul de învățare? și (2) poate „vaccinul” să fie generalizat și aplicat în alte domenii? De exemplu, deși este posibilă modelarea mesajelor de inoculare în funcție de un anumit subiect prin crearea de doze slabe de informații false, cum ar putea arăta un plan general de rezistență împotriva dezinformării? Pentru a răspunde la aceste întrebări, am lucrat împreună cu Dutch Media Collective „Bad News”, organizație care creează materiale educaționale inedite pentru combaterea dezinformării. Împreună, am lucrat pe baza cercetărilor noastre anterioare și am transpus descoperirile din laborator într-o experiență educațională interactivă: The Fake News Game. Am teoretizat că asumarea rolului unei persoane care încearcă activ să inducă în eroare va constitui un mod eficient de conferire a unei rezistențe generale împotriva dezinformării. Astfel, jocul pune utilizatorii în pielea unui producător de fake news. Inițial am creat o versiune pe hârtie a jocului (vezi Figura 9.4), în care elevii aleg un personaj specific, cum ar fi un adept al teoriilor conspirației sau un colporteur

de clickbait și structurează articole de știri într-o modalitate consecventă rolului.

Figura 9.4. Jocul de cărți Fake News (Roozenbeek și van der Linden, 2018)



La testul-pilot au participat 95 de elevi (cu vârste între 16–18 ani) de la un liceu olandez, contextul fiind constituit de articole fake news despre criza refugiaților sirieni (Roozenbeek și van der Linden, 2018). Clasele au fost repartizate aleatoriu într-un grup experimental și un grup de control. După ce au jucat 30 de minute, elevii din grupul experimental au evaluat ca fiind semnificativ mai puțin credibile știrile inventate, pe care le citeau pentru prima dată, despre criza refugiaților, în comparație cu grupul de control care nu a făcut decât să se uite la o înregistrare video ca parte a unui plan obișnuit de predare. Pentru a evalua dacă în acest proces a fost provocată perceperea amenințării, analiza atitudinilor prezente în răspunsurile la întrebări deschise a dezvăluit un afect negativ semnificativ mai intens în cadrul grupului experimental. Deși aceste rezultate au fost încurajatoare, studiul a avut o putere statistică relativ scăzută și elevii nu și-au schimbat neapărat atitudinile față de imigrare. Acest lucru este consecvent rezultatelor altor cercetări recente, care au descoperit că, deși intervențiile media pot reduce percepțiile eronate, asta nu se reflectă neapărat în schimbarea convingerilor politice (Nyhan et al., 2019). Însă e drept că aceste intervenții nu au scopul de a schimba convingerile politice, ci doar de a ajuta oamenii să observe tehnicile de dezinformare. Este esențial, în studiul nostru aplicat, că articolul fake news „atenuat” cu care elevii au fost „instruiți” a fost diferit de articolul cu care au fost testați — ceea ce oferă o dovadă preliminară potrivit căreia condițiile limită ale metaforei inoculării pot fi extinse.

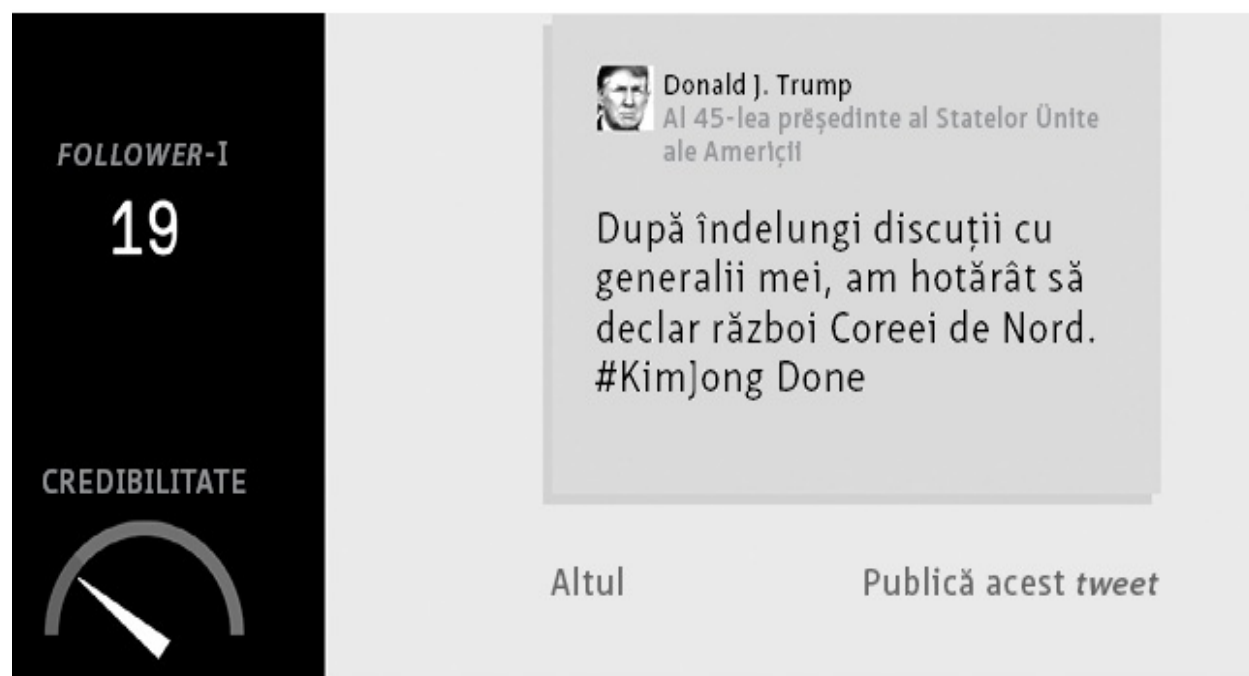
În mediul natural: jocul veștilor rele (studiul 3)

Pe baza acestor rezultate, am creat o versiune online a jocului Fake News (FastCompany, 2018), versiune care a câștigat mai multe premii. Am numit-o „Bad News” („Vești rele”) (Roozenbeek și van der Linden, 2019). Acest joc online simulează un motor social media (Twitter) și jucătorii trebuie să atragă

follower-i prin distribuirea de fake news. Interfața este prietenoasă (vezi Figura 9.5a); jucătorilor li se arată un text scurt sau o imagine (cum ar fi o memă sau un titlu) și pot reacționa la acestea prin mai multe moduri. Scorul este determinat în funcție de numărul de „follower-i” și de măsura „credibilității” (imaginea b). Scopul jocului este ca utilizatorii să adune cât mai mulți follower-i fără să-și piardă credibilitatea. Alegerea unei opțiuni conform unui producător „real” de informații false aduce mai mulți follower-i și o mai mare credibilitate. Însă, dacă își mint grosolan follower-ii sau dacă acționează conform bunelor practici jurnalistice, jocul fie scade numărul de follower-i, fie scade scorul credibilității.



(a)



(b)

Figura 9.5. Capturi de ecran din jocul *Bad News* (www.getbadnews.com)

Strategii de dezinformare

Conform metaforei inoculării, jocul expune utilizatorii la mai multe doze atenuate de dezinformare prin faptul de a le oferi posibilitatea de a genera activ propriul conținut. Cu toate acestea, prin opoziție cu inocularea împotriva unei teme anume, noi am emis ipoteza că ar putea fi posibilă „vaccinarea” împotriva înseși tacticilor care stau la baza producerii celor mai multe fake news (asemenea unui vaccin cu spectru larg). Întrucât este imposibilă acoperirea detaliată a tuturor aspectelor dezinformării într-un joc de 15 minute, am ales să ne concentrăm doar asupra celor mai comune strategii. Prin șase nivele de joc derivate pe baza teoriei, utilizatorii învață despre adoptarea unei identități false online, aplicarea limbajului afectiv, polarizarea grupurilor, teorii conspiraționiste, construirea de camere de ecou, discreditarea oponentilor, trolling și falsă amplificare. Aceste strategii sunt derivate parțial din raportul „Digital Hydra” al East Strategic Command (East StratCom) din cadrul NATO, care detaliază diferitele forme pe care le poate lua dezinformarea, precum și cercetările asupra strategiilor de inducere în eroare (Bertolin, Agarwal, Bandeli, Biteniece și Sedova, 2017). Secțiunea următoare oferă un rezumat al scenariilor și fundalului teoretic pentru fiecare mod de joc.

Identități false

Nu mai e dificilă construirea unui website și publicarea de conținut care pare într-un totu legitim. Dat fiind că aproape nu există obstacole în ce privește costul, aproape oricine poate deveni producător de conținut (Goga, Venkatadri și Gummati, 2015; Orita și Hada, 2009; Reznik, 2013). Scopul acestui mod de joc este să arate cât de simplu este acest proces și să

demonstreze că înfățișarea sau titlurile profesionale nu implică neapărat legitimitate. „Identitatea falsă” are două componente principale: (1) imitarea unei persoane sau a unei organizații reale prin simularea aparenței acesteia, de exemplu, prin folosirea unui username foarte asemănător, ca în farsa pe care a jucat-o cineva care în 2018 a pretins pe Twitter că era investitorul miliardar Warren Buffett (BBC News, 2018c), și (2) prezentarea ca website sau blog de știri legitim, însă fără normele și acreditările jurnalistice obișnuite.

În joc, utilizatorii mai întâi postează un tweet despre ceva ce îi frustrează, care poate fi orice, de la un guvern incompetent la teoria pământului plat. Asta aduce primii follower-i și apoi jocul explică modul în care funcționează numărătoarea follower-ilor și măsurătoarea credibilității. Utilizatorii simulează apoi că au contul oficial al lui Donald Trump (care declară război Coreei de Nord), al NASA (care anunță că un meteorit masiv e pe cale să lovească Pământul) sau al canalului TV Nickelodeon (care anunță anularea iminentă a desenelor animate SpongeBob SquarePants). Utilizatorilor li se arată apoi tweet-uri ale utilizatorilor de Twitter care au crezut aceste farse, după care jocul le solicită să facă lucrurile profesionist și să-și creeze propriul site de știri. Utilizatorii aleg un domeniu online, un titlu și un slogan.

Conținut afectiv

Conținutul afectiv este tipul de conținut care nu e neapărat „fals” sau „adevărat”, dar care mizează intenționat pe emoțiile fundamentale ale oamenilor, ca frica, furia sau empatia (Aday, 2010; Bakir și McStay, 2017; Gross și D’Ambrosio, 2004; Konijn, 2013). Scopul acestui mod de joc este să arate cum pot fi manipulate emoțiile de bază pentru agitarea follower-ilor, astfel încât conținutul să fie distribuit mai amplu.

Acesta este primul mod de joc în care utilizatorii produc conținut pentru site-ul fictiv de știri. Li se solicită să răsfoiască titluri ale știrilor pe care le pot publica pe site, având de ales între încălzirea globală și organisme modificate genetic (OMG-uri). Utilizatorilor li se cere apoi opinia despre subiectul ales.

Jocul le cere ca știrea să spună fie că tema aleasă va aduce apocalipsa (în cazul OMG-urilor), fie că nu e decât o farsă (în cazul încălzirii globale) ca modalitate de a strânge follower-i. Li se cere și să aleagă o abordare a subiectului: să atace oamenii de știință, să posteze o poveste încărcată afectiv sau să vorbească despre știință. Ultima opțiune determină un răspuns negativ, jucătorii fiind încurajați să apeleze la un tip reacționar de conținut pentru a-și agita follower-ii. Pot apoi fie să creeze o memă (un conținut media umoristic, de obicei o imagine sau un GIF, care se răspândește pe internet de la un individ la altul), fie să scrie un articol care le reflectă alegerea. Fiecare alegere implică numeroase opțiuni, dintre care una este întotdeauna negativă (fiindcă ratează esența). Unii follower-i vor reacționa la postarea pe Twitter într-un mod furios, încărcat afectiv, ceea ce reprezintă tocmai scopul jucătorului.

Polarizarea

Polarizarea implică intenția de a mări distanța dintre stânga și dreapta politică și de a-i îndepărta pe oameni de la centru (Bessi et al., 2016; Groenendyk, 2018; Melki și Pickering, 2014; Prior, 2013; Twenge, Honeycutt, Prislin și Sherman, 2016). Pentru a câștiga follower-i, site-urile noi de știri apelează adesea la polarizare ca modalitate de a-și crea o nișă în peisajul media online. Acest mod de joc acoperă și concepția de „falsă amplificare” sau ideea că nu e necesară crearea unui articol complet fals pentru a transmite un mesaj, ci, în loc de asta, pot fi amplificate nemulțumirile existente și pot fi făcute să pară mai comune decât sunt de fapt (Bertolin et al., 2017).

La începerea acestui nivel, utilizatorii sunt întrebați dacă doresc să publice o informație falsă sau ceva real. Dacă aleg informația „falsă”, li se spune că nu e întotdeauna necesar să inventeze fake news pentru a obține vizualizări, ci că pot și să găsească o informație adevărată și să-i amplifice proporțiile. Utilizatorii pot apoi să mărească distanța dintre stânga și dreapta politică alegând o știre dintre trei articole locale, raportate de utilizatori anonimi de Twitter: o scurgere de substanțe chimice, un scandal de mită dintr-un orașel sau arestarea brutală a unui infractor. Mai întâi, utilizatorii aleg o țintă: în

două cazuri pot să atace fie marile corporații, fie guvernul, și în celălalt caz pot să atace fie poliția, fie infractorii violenți. Ei pot apoi să expună informația din perspectiva aleasă pe contul de Twitter al site-ului de știri, dar asta eșuează. Li se cere să amplifice povestea prin publicarea unui articol sau prin postarea unei meme. Asta le aduce mai mulți follower-i, pe măsură ce oamenii încep să fie interesați de poveste. Apoi jocul întreabă utilizatorii dacă doresc să achiziționeze boți Twitter care pot amplifica articolul. Dacă refuză în repetate rânduri, jocul se termină, dar, dacă acceptă, obțin 4 000 de follower-i boți. Li se arată exemple de astfel de boți care amplifică povestea aleasă. Alegerile utilizatorilor determină dacă polarizarea tinde spre stânga (centrându-se pe marile corporații sau pe brutalitatea poliției) sau spre dreapta (centrându-se asupra guvernului sau infraționalității). Lecția esențială este că nu contează ce direcție aleg: scopul este pur și simplu polarizarea.

Conspirația

Teoriile conspirației reprezintă o parte esențială a site-urilor de știri de la periferia online. Conspiraționismul poate fi definit drept convingere că evenimente neexplicate sunt orchestrate de un grup sau de o organizație secretă, cu intenții sinistre (Goertzel, 1944; Lewandowsky, Gignac și Oberauer, 2013; van der Linden, 2015).

Pe acest nivel, utilizatorii sunt mai întâi încurajați să publice pe site-ul lor de știri o teorie nouă și interesantă. Însă, dat fiind că toate opțiunile sunt exagerat de ridicole (e.g. școlile publice nu mai predau scrisul cu litere cursive, astfel încât oamenii să nu mai poată citi Manifestul Comunist), teoria lor este prea îndepărtată de realitate pentru a fi credibilă. Unii follower-i îi atacă pentru a fi publicat o teorie atât de stranie. Pentru a-și salva credibilitatea, utilizatorii caută apoi o conspirație mai credibilă. Pot să aleagă fie Agenda 21, un acord al Națiunilor Unite asupra dezvoltării sustenabile, fie așa-numita conspirație a vaccinurilor (ideea că Organizația Mondială a Sănătății folosește vaccinuri pentru a îndoctrina oamenii). Utilizatorii obțin puncte dacă pot să pună la îndoială perspectiva oficială și pun întrebări care îi fac pe oameni să tindă într-o direcție conspiraționistă; pierd puncte dacă

depășesc limita și publică un conținut considerat prea bizar. Follower-ii reacționează mai eficace de data asta și utilizatorii sunt încurajați să publice un articol serios despre subiectul ales. Dacă se descurcă bine, ajung să aibă follower-i care îi urmăresc cu loialitate de sectanți, cu din ce în ce mai multă încredere în site-ul lor de știri și care sunt din ce în ce mai sceptici în privința mass-media.

Discreditarea adversarilor

Când site-urile de știri care induc în eroare sunt acuzate de jurnalism prost, pot să distragă atenția de la acuzație prin atacarea sursei criticii („voi sunteți fake news!“, vezi van der Linden, Panagopoulos și Roozenbeek, 2020) sau prin negarea faptului că problema există (A’Beckett, 2013; Lischka, 2017). Pe acest nivel, utilizatorii se confruntă cu o persoană care verifică faptele și care demască teoria conspirației de la nivelul anterior. Li se oferă trei opțiuni: să-și ceară scuze, să nu facă nimic sau să se răzbune. Prima opțiune scade scorul și moderatorul jocului explică faptul că prezentarea de scuze nu este niciodată o idee bună. Dacă aleg să nu facă nimic, jocul prezintă un răspuns din partea unui follower al site-ului care îi întreabă de ce nu răspund. În cele din urmă, toate trei opțiunile duc în același punct în care utilizatorii trebuie să aleagă dacă să nege acuzațiile sau să atace. Negarea vehementă sau atacul la persoană determină reacții pozitive din partea follower-ilor și reputația jucătorului rămâne intactă.

Trolling

Trolling este un termen din sfera pescuitului, referindu-se la procesul de a trage lent o momeală înapoi în barcă. Pe internet înseamnă provocarea intenționată a unei reacții prin folosirea unei momeli (Griffiths, 2014; McCosker, 2014; Thacker și Griffiths, 2012).

Pe acest nivel, utilizatorii pun laolaltă tehnici pe care le-au învățat pe celelalte cinci nivele. De data asta pot să aleagă un singur subiect. La început li se cere să vorbească despre una din trei teme (25 cele mai romantice orașe din Europa; prăbușirea unui avion de pasageri; o specie nou descoperită de stea-de-mare), dintre care doar a doua duce la etapa următoare. Alegerea uneia dintre celelalte două atrage critica moderatorului jocului. Apoi utilizatorilor li se oferă două opțiuni: fie aduc omagiu victimelor accidentului aviatic, fie încep să arunce îndoieli asupra cauzei. Prima opțiune determină o reacție a follower-ilor care întreabă de ce povestea nu e investigată mai în detaliu. Ambele opțiuni fac utilizatorul ca în cele din urmă să întrebe dacă nu cumva accidentul a fost o mușamalizare. Datorită nivelului ridicat de credibilitate și a numărului mare de follower-i, postarea lor atrage atenția altor site-uri de știri și povestea escaladează. Jucătorii pot apoi să pună paie pe foc fie prefăcându-se că sunt un membru al familiei uneia dintre victime, fie falsificând imaginile, astfel încât să pară că aduc dovezi ale unei mușamalizări. Ambele variante duc apoi la reacții afective și mai intense, și acum încep să fie interesate și posturile mass-media. Jucătorilor li se spune să crească presiunea, fie prin continuarea discreditării investigației, fie prin folosirea unei alte armate de boți Twitter care să răspândească hashtag-ul #InvestigateNow (Investigați acum). În funcție de alegerea inițială între imitarea unui membru al familiei uneia dintre victime sau falsificarea unei imagini, jucătorii pot apoi să dea ultima lovitură inventând un alt articol de știri despre prăbușirea avionului. Comisia pentru Dezastre Aviatice, agenția (fictivă) responsabilă de investigație, răspunde apoi pe Twitter la această controversă inventată. Jucătorii atacă acest răspuns fie cerând demisia directorului Comisiei, fie folosindu-se din nou de armata de boți. Jocul se încheie cu demisia directorului, din cauza modului în care a coordonat investigația.

Lansare și rezultatele anchetei

Jocul a apărut în februarie 2018, iar lansarea acestuia a fost mediatizată

internațional, atât online, cât și prin mijloacele tradiționale (BBC, 2017; Reuters, 2018; CNN, 2019). Jocul includea participarea voluntară la o anchetă de tip pre-post, care evalua capacitatea oamenilor de a recunoaște tehnicile de dezinformare. Pe parcursul a două luni, sute de mii de oameni au jucat jocul (majoritatea, din Marea Britanie și SUA). În total, aproximativ $N = 15\,000$ de persoane au optat să participe la cercetarea științifică și au completat toate testele anchetei (pre-post). Fiecare întrebare apărea sub forma unui tweet inventat care reprezenta o anumită strategie de dezinformare. Specific, într-o evaluare inițială am testat capacitatea participanților de a recunoaște falsificarea identității (prin intermediul unui cont care pretindea că aparține HBO și spunea „Sezonul 8 al Game of Thrones va fi amânat din cauza neînțelegerilor salariale”), conspirația (un tweet care spunea „Rata de schimb Bitcoin este manipulată de un grup restrâns de bancheri bogați”), discreditarea adversarilor (un alt tweet care afirma că „Mass-media au fost prinse cu atâtea minciuni, că nu mai pot fi surse de încredere”) și polarizarea (le-am arătat participanților titluri inventate care spuneau fie „Un nou studiu arată că stânga politică minte mai mult decât dreapta” sau, invers, „Un nou studiu arată că dreapta politică minte mai mult decât stânga”). Participanților li s-a cerut să evalueze credibilitatea fiecăruia dintre aceste tweet-uri pe o scală de la 1 la 7, plus două tweet-uri de control „reale”, în care nu apărea nicio strategie de dezinformare (e.g. #Brexit, ieșirea Marii Britanii din Uniunea Europeană, va avea loc oficial în 2019). Am folosit titluri fictive inspirate de articole fake news „reale” din două motive: (1) pentru a putea izola tehnicile specifice de dezinformare și (2) pentru a evita confundarea familiarității cu un conținut fake news „real” (Roozenbeek și van der Linden, 2019). Un exemplu de testare este oferit în figura 9.6.

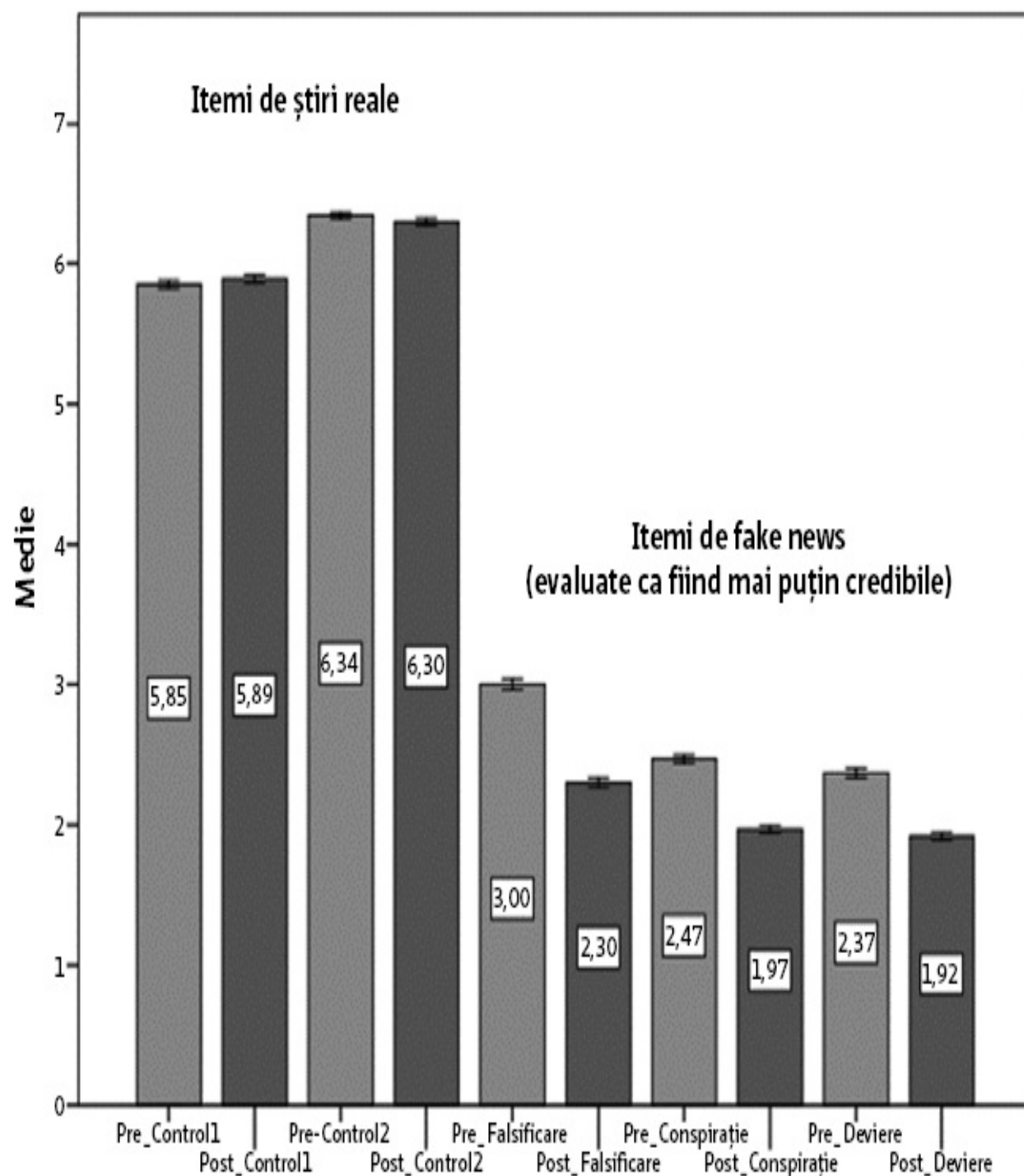


Figura 9.6. Ancheta realizată prin jocul *Bad News*

Notă: Imaginea de sus ilustrează modul în care o tehnică [identitatea falsă] este folosită în joc, iar imaginea de jos arată modul în care aceeași tehnică este folosită într-un exemplu diferit prin care participanții au fost evaluați înainte și după ce au jucat.

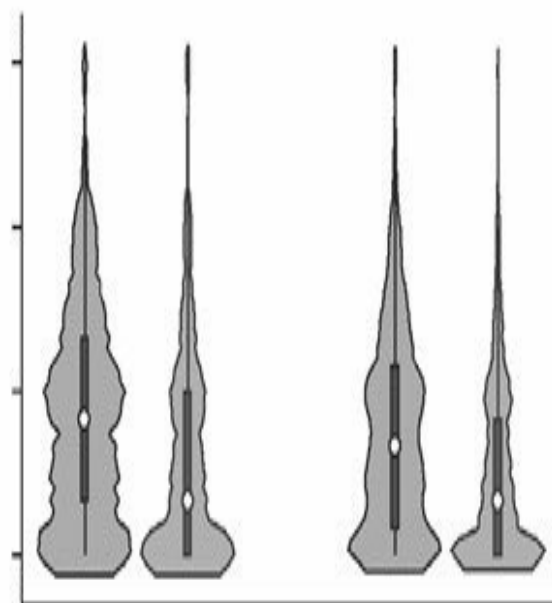
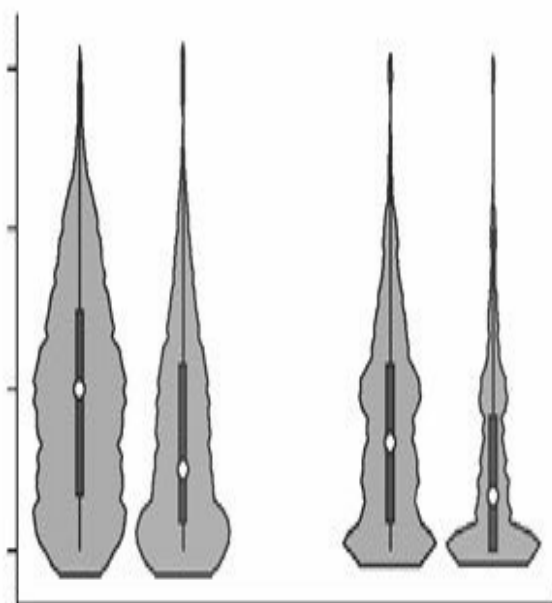
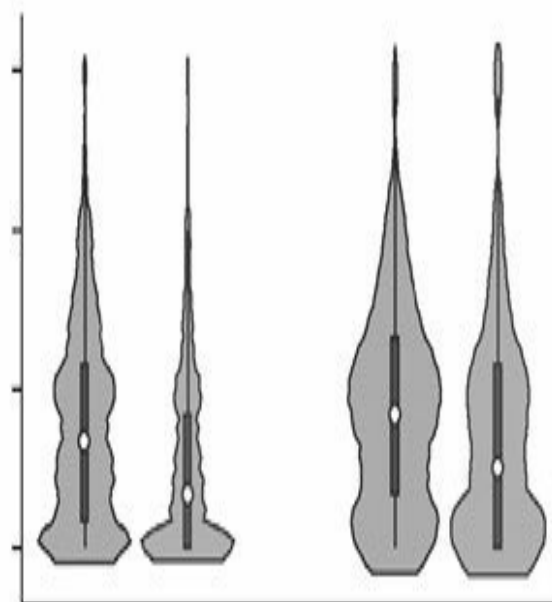
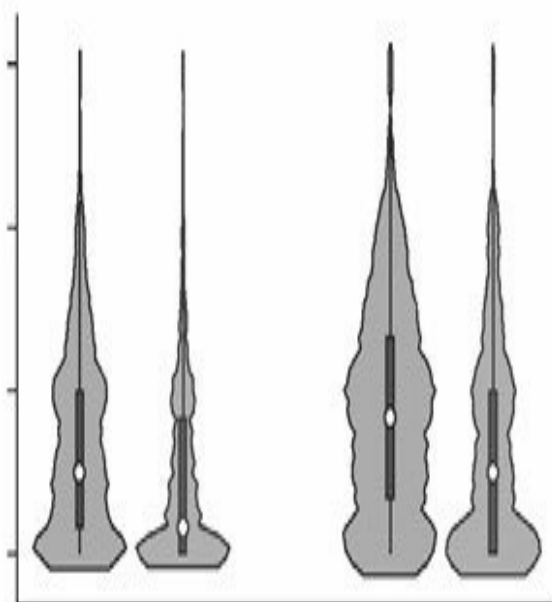
Rezultatele sunt expuse în Figurile 9.7 și 9.8. În figura 9.7, graficul-bloc din stânga al fiecărei perechi este răspunsul oferit de participanți înainte de a juca (gri-deschis), iar graficul-bloc din dreapta este răspunsul oferit după ce au jucat (gri-închis). Un motiv inițial de îngrijorare era posibilitatea ca jocul pur și simplu să îi facă pe utilizatori să fie mai sceptici față de media în general. Nu asta am descoperit. „Control_1” și „Control_2” reprezintă întrebările de control. Ne-am așteptat ca participanții să le evalueze ca fiind credibile, și înainte, și după ce au jucat. Fapt confirmat de rezultate: în ambele cazuri, măsurătorile pre și post sunt aproape identice și indică un nivel ridicat de credibilitate. Celelalte întrebări reprezintă tehnicile utilizate frecvent în dezinformare și care apar și în joc: identitatea falsă, conspirația și discreditarea adversarilor. După cum se vede în figura 9.7, participanții au dat scoruri semnificativ mai mici la întrebările „fake news” după ce au jucat, ceea ce indică un efect semnificativ de inoculare la falsificarea identității ($d = 0,36$), conspirație ($d = 0,35$), deviere ($d = 0,30$) și polarizare ($d = 0,16$). Aceste mărimi ale efectului corespund rezultatelor obținute de cercetările asupra rezistenței la persuasiune (Walter și Murphy, 2018) și pot fi considerate semnificative, mai ales dacă sunt extinse la nivelul populației (Funder și Ozer, 2019). Este important faptul că efectele de învățare nu au prezentat diferențe semnificative în funcție de ideologia politică, de vârstă, de gen sau de nivel de educație (Figura 9.8) și au fost mai mari în cazul celor care s-au dovedit mai susceptibili la pre-test, ceea ce amplifică potențialul imunizării cu spectru larg. Cu alte cuvinte, prin inoculare activă împotriva strategiilor subiacente fake news — prin expunere la conținuturi atenuate — se poate conferi o rezistență la un spectru larg de metode de dezinformare.

Figura 9.7. Rezultate obținute din jocul *Bad News*



Notă: Blocurile pereche reprezintă evaluările pe care utilizatorii le-au făcut înainte (gri-deschis) și după (gri-închis) ce au jucat. Participanții au evaluat credibilitatea itemilor reali (de control) și *fake news* pe o scală de la 1 la 7.

Figura 9.8. Grafic vioară (violin plot) adaptat după Roozenbeek și van der Linden (2019) care reprezintă media pre-post evaluărilor credibilității fake news (agregate) în funcție de ideologia politică (A), vârstă (B), educație (C) și a unei măsurători a reflecției cognitive cu un 1 item (D)



A

C

7

5

3

1

7

5

3

1

Liberal (stânga politică)

Nivel redus de educație

Pre-test

Pre-test

Conservator (dreapta politică)

Nivel înalt de educație

Post-test

Post-test

B

D

7

5

3

1

7

5

Intuitiv (CRT)

Tineri (18–49)

Pre-test

Pre-test

Reflexiv (CRT)

Maturi (50+)

Post-test

Post-test

Limite și concluzie

În acest capitol am examinat motivele pentru care dezinformarea reprezintă o problemă și am investigat tipurile de soluții care pot fi folosite pentru combaterea acesteia. Din perspectiva noastră, soluția constă în principal în faptul ca indivizilor să li se ofere instrumente psihologice și comportamentale bazate pe dovezi. Am susținut că este deosebit de importantă concentrarea asupra împiedicării faptului ca dezinformarea să ajungă virală. De fapt, răspândirea fake news poate fi modelată în mare parte similar răspândirii unui virus contagios. Astfel, teoria inoculării reprezintă un cadru intuitiv care ne

poate ajuta să dezvoltăm metode de imunizare cu spectru larg împotriva fake news.

Prin intermediul a trei studii am arătat că metafora inoculării poate fi extinsă din domeniul truismelor culturale în sfera chestiunilor controversate și am expus modul în care putem trece de la vaccinuri pentru cazuri specifice de fake news la o abordare cu un spectru larg de inoculare a oamenilor împotriva strategiilor care stau la baza dezinformării. Desigur, aceste studii nu sunt lipsite de limitări. În primul rând, deși eșantioanele au fost diversificate, au fost alcătuite fie din studenți, fie din persoane recrutate prin Mechanical Turk, fie din indivizi din populația generală care s-au oferit ei înșiși voluntari. Astfel, aceste rezultate nu sunt reprezentative pentru populația generală. În plus, studiul 3 a folosit evaluări cu un singur item și nu a avut un grup de control cu eșantionare aleatorie. Așadar, efectele cauzale trebuie interpretate cu precauție. Mai mult, deși metodele de inoculare s-au dovedit eficiente pe întreg spectrul ideologiilor politice, utilizatorii jocului au putut să modeleze scenariile în modalități relativ consecvente propriei ideologii. Cu toate astea, documentarea procesului tranzițional al modului în care putem trece de la un rezultat teoretic în laborator la o intervenție în lumea reală, noi subliniem potențialul educațional al unei intervenții noi, interactive, bazate pe joc. Valoarea intervențiilor științifice se poate vedea în multe dintre aplicațiile sale. De exemplu, în parteneriat cu guvernul Marii Britanii, am tradus jocul în 12 limbi, ceea ce a făcut posibilă evaluarea la scară largă a eficienței jocului în cadrul mai multor culturi (Roozenbeek, van der Linden și Nygre, 2020). Am lucrat în parteneriat și cu WhatsApp pentru a crea o versiune specială a jocului pentru a ajuta oamenii să se inoculeze împotriva răspândirii de fake news pe platformele de mesagerie instantă. Lucrăm împreună cu guvernul SUA și Behavioral Insights Team din Marea Britanie și Liban pentru a realiza anchete randomizate la scară mai mare, astfel încât să continuăm evaluarea și îmbunătățirea acestei intervenții. Pe scurt, dacă noua știință a predemascării este atât de eficientă pe cât pare, viitoare cercetări sociale și comportamentale vor putea contribui la cultivarea imunității sociale împotriva fake news.

Notă

1.Cercetările ulterioare nu au avut întotdeauna aceleași rezultate (vezi Banas și Rains, 2010), poate din cauza unei încărcări cognitive mai mari, asociate faptului că participanții trebuiau să genereze propriile contraargumente.

Referințe

A'Beckett, L. (2013). Strategies to discredit opponents: Russian representations of events in countries of the former Soviet Union. *Psychology of Language and Communication*, 17 (2). <https://doi.org/10.2478/plc-2013-0009>

Aday, S. (2010). Leading the charge: Media, elites, and the use of emotion in stimulating rally effects in wartime. *Journal of Communication*, 60 (3), 440–465. <https://doi.org/10.1111/j.1460-2466.2010.01489.x>

Adl-Tabatabai, S. (2016). Australia to forcibly vaccinate citizens via Chemtrails. Accesat la 8 august 2018, <https://yournewswire.com/australia-to-forcibly-vaccinatecitizens-via-chemtrails/>

Bakir, V. și McStay, A. (2017). Fake news and the economy of emotions: Problems, causes, solutions. *Digital Journalism*, 1–22. <https://doi.org/10.1080/21670811.2017.1345645>

Banas, J.A. și Rains, S.A. (2010). A meta-analysis of research on inoculation theory. *Communication Monographs*, 77 (3), 281–311. <https://doi.org/10.1080/03637751003758193>

Barthel, M., Mitchell, A. și Holcomb, J. (decembrie 2016). Many Americans believe fake news is sowing confusion. Pew Research Center. Accesat la www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/

Baxter, D. (2018). Melania Trump bans white house staff from taking flu shot. Accesat la 8 august 2018, <http://archive.is/F0vor>

BBC. (2017). Cambridge scientists consider fake news „vaccine“. Accesat la 29 august 2017, www.bbc.co.uk/news/uk-38714404

BBC News. (2018a, 19 iulie). How WhatsApp helped turn an Indian village into a lynch mob. Accesat la www.bbc.co.uk/news/world-asia-india-44856910

BBC News. (2018b). Government announces anti-fake news unit. Accesat la 8 august 2018, from www.bbc.co.uk/news/uk-politics-42791218

BBC News. (2018c, 28 august). A fake billionaire is fooling people on Twitter. Accesat la www.bbc.co.uk/news/world-us-canada-45331781

Bertolin, G., Agarwal, N., Bandeli, K., Biteniece, N. și Sedova, K. (2017). Digital Hydra: Security implications of false information online. Accesat la www.stratcomcoe.org/digital-hydra-security-implications-false-information-online

Bessi, A., Zollo, F., Del Vicario, M., Puliga, M., Scala, A., Caldarelli, G., ... Quattrociocchi, W. (2016). Users polarization on Facebook and Youtube. PLoS One, 11 (8), 1–24. <https://doi.org/10.1371/journal.pone.0159641>

Bremner, C. (2018). France aims to ban fake news at election time. Accesat la www.thetimes.co.uk/article/france-aims-to-ban-fake-news-at-election-time-jwspzjx83

Calfas, J. (2017). Google is changing its search algorithm to combat fake news. Accesat la 8 august 2018, <http://fortune.com/2017/04/25/google-search-algorithmfake-news/>

Channel 4. (2017). C4 study reveals only 4% surveyed can identify true or fake news. Accesat la 29 august 2017, www.channel4.com/info/press/news/c4-studyreveals-only-4-surveyed-can-identify-true-or-fake-news

CNN. (2019). Researcher have created a „vaccine“ for fake news: It’s a game. Accesat la https://edition.cnn.com/2019/07/04media/fake-news-game-vaccine/index.html?utm_medium=social&utm_content=2019-07-04T13%3A46%3A03&utm_source=twCNN&utm_term=link

Compton, J. (2013). Inoculation theory. În J.P. Dillard și L. Shen (coord.), *The SAGE handbook of persuasion: Developments in theory and practice* (ed. a II-a, pp. 220–236). Thousand Oaks: Sage Publications.
<https://doi.org/10.4135/9781452218410>

Compton, J. (2019). Prophylactic versus therapeutic inoculation treatments for resistance to influence. *Communication Theory*, qtz004.

Compton, J., Jackson, B. și Dimmock, J.A. (2016). Persuading others to avoid persuasion: Inoculation theory and resistant health attitudes. *Frontiers in Psychology*, 7, 122.

Cook, J., Lewandowsky, S. și Ecker, U.K. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PloS One*, 12 (5), e0175799.
<https://doi.org/10.1371/journal.pone.0175799>

Cook, J., Oreskes, N., Doran, P.T., Anderegg, W.R.L., Verheggen, B., Maibach, E.W., ... Rice, K. (2016). Consensus on consensus: A synthesis of consensus estimates on human-caused global warming. *Environmental Research Letters*, 11 (4), 48002. Accesat la <http://stacks.iop.org/1748-9326/11/i=4/a=048002>

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., ... Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113 (3), 554–559.
<https://doi.org/10.1073/pnas.1517441113>

Eagly, A.H. și Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jovanovich.

Elgin, B. și Wang, S. (2018). Facebook’s battle against fake news Notches an Uneven scorecard. Accesat la 8 august 8, 2018,

www.bloomberg.com/news/articles/2018-04-24/facebook-s-battle-against-fake-news-notches-an-uneven-scorecard

Eurobarometer (2018). Final results of the Eurobarometer on fake news and online disinformation. Directorate-General CONNECT, European Commission. Accesat la <https://ec.europa.eu/digital-single-market/en/news/final-results-eurobarometerfake-news-and-online-disinformation>

European Commission. (2018). Final report of the high level expert group on fake news and online disinformation. Accesat la <https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation>

FastCompany. (2018). Innovation by design awards: University of Cambridge/DROG. Accesat la www.fastcompany.com/innovation-by-design/2018/company/university-of-cambridge-drog

Ferrara, E. (2017). Disinformation and social bot operations in the run up to the 2017 French presidential election. *First Monday*, 22 (8). Accesat la https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2995809

Funder, D.C. și Ozer, D.J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, 2 (2), 156–168.

Goertzel, T. (1994). Belief in conspiracy theories. *Political Psychology*, 15 (4), 731–742. <https://doi.org/10.2307/3791630>

Goga, O., Venkatadri, G. și Gummadi, K.P. (2015). The doppelgänger bot attack: Exploring identity impersonation in online social networks. În *Proceedings of the 2015 internet measurement conference* (pp. 141–153). New York, NY, USA: ACM. <https://doi.org/10.1145/2815675.2815699>

Griffiths, M.D. (2014). Adolescent trolling in online environments: A brief overview. *Education and Health*, 32 (3), 85–87. Accesat la <http://irep.ntu.ac.uk/id/eprint/25950/>

Groenendyk, E. (2018). Competing motives in a polarized electorate:

Political responsiveness, identity defensiveness, and the rise of partisan antipathy. *Political Psychology*, 39, 159–171.
<https://doi.org/10.1111/pops.12481>

Gross, K. și D'Ambrosio, L. (2004). Framing emotional response. *Political Psychology*, 25 (1), 1–29. Accesat la www.jstor.org/stable/3792521

Hovland, C.I., Lumsdaine, A.A. și Sheffield, F.D. (1949). Experiments on mass communication. (Studies in social psychology in World War II) (Vol. 3). Princeton, NJ: Princeton University Press.

Huddy, L., Sears, D.O. și Levy, J.S. (2013). *The Oxford handbook of political psychology*. Oxford: Oxford University Press.

Konijn, E.A. (2013). The role of emotion in media use and effects. *The Oxford Handbook of Media Psychology*, 186–211.

Kragh, M. și Åsberg, S. (2017). Russia's strategy for influence through public diplomacy and active measures: The Swedish case. *Journal of Strategic Studies*, 40 (6), 773–816.
<https://doi.org/10.1080/01402390.2016.1273830>

Kurtzleben, D. (2016). Do fact checks matter? Accesat la 8 2018, www.npr.org/2016/09/27/495233627/do-fact-checks-matter

Lazer, D.M.J., Baum, M.A., Benkler, Y., Berinsky, A.J., Greenhill, K.M., Menczer, F., ... Zittrain, J.L. (2018). The science of fake news. *Science*, 359 (6380), 1094–1096. <https://doi.org/10.1126/science.aao2998>

Le Pen, M. (19 martie 2018). Postare pe Twitter. Accesată la 30 august 2018, https://twitter.com/mlp_officiel/status/975808872802856960?lang=en

Lewandowsky, S., Ecker, U.K.H. și Cook, J. (2017). Beyond misinformation: Understanding and coping with the „post-truth“ era. *Journal of Applied Research in Memory and Cognition*, 6 (4), 353–369.
<https://doi.org/10.1016/j.jarmac.2017.07.008>

Lewandowsky, S., Ecker, U.K.H., Seifert, C.M., Schwarz, N. și Cook, J.

(2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13 (3), 106–131. <https://doi.org/10.1177/1529100612451018>

Lewandowsky, S., Gignac, G.E. și Oberauer, K. (2013). The role of conspiracist ideation and worldviews in predicting rejection of science. *PLoS One*, 8 (10), 1–11. <https://doi.org/10.1371/journal.pone.0075637>

Lischka, J.A. (2017). A badge of honor? How the New York Times discredits President Trump's fake news accusations. *Journalism Studies*, 1–18. <https://doi.org/10.1080/1461670X.2017.1375385>

McCosker, A. (2014). Trolling as provocation: YouTube's agonistic publics. *Convergence*, 20 (2), 201–217. <https://doi.org/10.1177/1354856513501413>

McGuire, W.J. (1970). A vaccine for brainwash. *Psychology Today*, 3 (9), 36–64.

McGuire, W.J. și Papageorgis, D. (1961). The relative efficacy of various types of prior belief-defense in producing immunity against persuasion. *Journal of Abnormal and Social Psychology*, 62, 327–337.

McGuire, W.J. și Papageorgis, D. (1962). Effectiveness of forewarning in developing resistance to persuasion. *Public Opinion Quarterly*, 26 (1), 24–34. <https://doi.org/10.1086/267068>

Melki, M. și Pickering, A. (2014). Ideological polarization and the media. *Economics Letters*, 125 (1), 36–39. <https://doi.org/10.1016/j.econlet.2014.08.008>

Nyhan, B., Porter, E., Reifler, J. și Wood, T.J. (2019). Taking fact-checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability. *Political Behavior*, 1–22.

Orita, A. și Hada, H. (2009). Is that really you? An approach to assure identity without revealing real-name online. In *Proceedings of the 5th ACM workshop on digital identity management* (pp. 17–20). New York, NY, USA: ACM. <https://doi.org/10.1145/1655028.1655034>

Papageorgis, D. și McGuire, W.J. (1961). The generality of immunity to persuasion produced by pre-exposure to weakened counterarguments. *Journal of Abnormal and Social Psychology*, 62, 475–481.

Patriot United. (2018). Muslim doctor refuses to save a Christian man's life on flight from New York to Las Vegas. Accesat la 8 august 2018, <https://web.archive.org/web/20180108183412/http://patriotunited.club/2018/01/doctor-refusesto-save-a-christian-mans-life-on-flight-from-new-york-to-las-vegas/>

Pfau, M., Park, D., Holbert, R.L. și Cho, J. (2001). The effects of party-and PAC-sponsored issue advertising and the potential of inoculation to combat its impact on the democratic process. *American Behavioral Scientist*, 44 (12), 2379–2397.

Pieters, J. (2018). Dutch politicians want EU anti-fake news watchdog scrapped. Accesat la 8 august 8, 2018, <https://nltimes.nl/2018/03/06/dutch-politicianswant-eu-anti-fake-news-watchdog-scrapped>

Poland, G.A. și Spier, R. (2010). Fear, misinformation, and innumerates: How the Wakefield paper, the press, and advocacy groups damaged the public health. *Vaccine*, 28 (12), 2361–2362.

Prior, M. (2013). Media and political polarization. *Annual Review of Political Science*, 16 (1), 101–127. <https://doi.org/10.1146/annurev-polisci-100711-135242>

Readfearn, G. (2016). Revealed: Most popular climate story on social media told half a million people the science was a Hoax. Accesat la 8 august 2018, www.desmogblog.com/2016/11/29/revealed-most-popular-climate-story-social-mediataold-half-million-people-science-was-hoax

Reuters. (2018). Fake news „vaccine“ teaches you to spot disinformation. Accesat la 15 ianuarie 2019, <https://uk.reuters.com/video/2018/03/20/fake-news-vaccineteaches-you-to-spot-di?videoId=410596269>

Reznik, M. (2013). Identity theft on social networking sites: Developing issues of internet impersonation. *Touro Law Review*, 29 (2), 12.

Roozenbeek, J. și van der Linden, S. (2018). The fake news game: Actively inoculating against the risk of misinformation. *Journal of Risk Research*, 22 (5), 570–580. <https://doi.org/10.1080/13669877.2018.1443491>

Roozenbeek, J., și van der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. *Nature Palgrave Communications*, 5 (65). <https://doi.org/10.1057/s41599-019-027>

Roozenbeek, J., van der Linden, S., și Nygren, T. (2020). Prebunking interventions based on „inoculation“ theory can reduce susceptibility to misinformation across cultures. *Harvard Kennedy School Misinformation Review*, 1 (2). <https://doi.org/10.37016//mr-2020-008>

Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.C., Flammini, A. și Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature Communications*, 9 (1),4787.

Snopes. (2018). Did „Muslim migrants“ attack a Catholic Church during mass in France? Accesat la 30 august 2018, www.snopes.com/fact-check/muslim-migrantsattack-catholic-church-mass-france/

Tandoc, E.C., Jr., Lim, Z.W. și Ling, R. (2018). Defining „fake news“ a typology of scholarly definitions. *Digital Journalism*, 6 (2), 137–153.

Thacker, S. și Griffiths, M.D. (2012). An exploratory study of trolling in online video gaming. *International Journal of Cyber Behavior, Psychology and Learning*, 2 (4). <https://doi.org/10.4018/ijcbpl.2012100102>

Twenge, J.M., Honeycutt, N., Prislin, R. și Sherman, R.A. (2016). More polarized but more independent: Political party identification and ideological self-categorization among U.S. adults, college students, and late adolescents, 1970–2015. *Personality and Social Psychology Bulletin*, 42 (10), 1364–1383. <https://doi.org/10.1177/0146167216660058>

van der Linden, S. (2015). The conspiracy-effect: Exposure to conspiracy theories (about global warming) decreases pro-social behavior and science acceptance. *Personality and Individual Differences*, 87, 171–173. Accesat la www.sciencedirect.com/science/article/pii/S0191886915005024

van der Linden, S. (2017). Beating the hell out of fake news. *Ethical Record: Proceedings of the Conway Hall Ethical Society*, 122 (6), 4–7. Accesat la <https://ssrn.com/abstract=3089590>

van der Linden, S., Leiserowitz, A., Rosenthal, S., și Maibach, E. (2017). Inoculating the public against misinformation about climate change. *Global Challenges*, 1 (2), 1600008. <https://doi.org/10.1002/gch2.201600008>

van der Linden, S., Maibach, E., Cook, J., Leiserowitz, A. și Lewandowsky, S. (2017). Inoculating against misinformation. *Science*, 358 (6367), 1141–1142. <https://doi.org/10.1126/science.aar4533>

van der Linden, S., Panagopoulos, C. și Roozenbeek, J. (2020). You are fake news: political bias in perceptions of fake news. *Media, Culture & Society*, 42 (3), 460–470.

van der Linden, S., Roozenbeek, J., Oosterwoud, R., Compton, J. și Lewandowsky, S. (2018). The science of prebunking: Inoculating the public against fake news. Written evidence submitted to the Parliamentary Inquiry on Fake News. House of Commons, UK Parliament, Londra: Marea Britanie. Accesat la <http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedoc/culture-media-and-sportcommittee/fake-news/written/79482.html>

Varol, O., Ferrara, E., Davis, C.A., Menczer, F. și Flammini, A. (2017). Online human-bot interactions: Detection, estimation, and characterization. *Proceedings of the Eleventh International AAAI Conference on Web and Social Media*. Accesat la www.aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/viewPaper/15587

Vosoughi, S., Roy, D. și Aral, S. (2018). The spread of true and false news online. *Science*, 359 (6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>

Wakabayashi, D. (2017). As Google fights fake news, voices on the margins raise alarm. Accesat la 15 february 2018, www.nytimes.com/2017/09/26/technology/google-search-bias-claims.html

Wakefield, J. (2017). Facebook's fake news experiment backfires. Accesat la 8 august 2018, www.bbc.co.uk/news/technology-41900877

Walter, N. și Murphy, S.T. (2018). How to unring the bell: A meta-analytic approach to correction of misinformation. *Communication Monographs*, 85 (3), 423–441.

***** Aluzie la Testul de personalitate Big Five, construit pe baza a cinci trăsături majore de personalitate: extraversie, agreabilitate, conștiinciozitate, deschidere și nevrotism. (N.r.)

***** Titlu provocator sau senzationalist care tentează cititorii să deschidă link-ul către articol (să dea click), folosit ca tactică editorială de a crește numărul de vizualizări pe pagină și, implicit, venitul din publicitate. (N.r.)

***** Incident fictiv de terorism invocat la începutul anului 2017 de consilierul prezidențial american Kellyanne Conway ca justificare a interdicției de imigrare impuse de Donald Trump pentru șapte state majoritar musulmane. Identificat rapid ca fals, Conway s-a corectat, spunând că de fapt s-a referit la arestarea în 2011 a doi susținători al-Qaeda în Bowling Green, Kentucky. Însă declarația a devenit virală și a provocat o reacție amplă de parodiare pe Twitter și Facebook. Au fost organizate priveghiuri și o campanie de donații pentru supraviețuitori. (N.t.)

***** Russian Bots sau web brigades este numele dat comentatorilor politici anonimi de pe internet, ce se consideră a fi sponsorizați de statul rus. (N.r.)

***** Limba populației amerindiene Hopi, din Arizona, SUA. (N.t.)

**PARTEA a III-a. Procese motivaționale implicate în acceptarea,
distribuirea și corectarea dezinformării**

Capitolul 10. Voi, cu fake news, noi, cu faptele

Motivațiile bazate pe identitate modelează ce credem, ce distribuim și ce acceptăm

Daphna Oyserman și Andrew Dawson

Introducere

La 23 iunie 2016, cetățenii britanici s-au dus să voteze, sau, mai degrabă, șapte din zece englezi s-au dus să voteze; ceilalți s-au abținut (The Guardian, 2016). A fost surprinzător faptul că prezența la vot nu a fost de 100%, întrucât miza era rămânerea Marii Britanii (Anglia, Irlanda de Nord, Scoția și Țara Galilor) în Uniunea Europeană (UE), așa cum a fost din 1973. UE a fost construită pe premisa că membrii ei sunt mai în siguranță, mai puternici și mai liberi împreună — e mai puțin probabil să intre în război; economia fiecăreia e mai prosperă; cetățenii își pot alege propria cale. O întreagă generație de britanici a crescut într-un mediu în care Londra a fost centrul financiar al UE (Brush și Weber, 2019), universitățile britanice primind finanțări din partea UE (UK Research and Innovation, 2019) și produsele britanice trecând liber prin UE, cel mai mare partener de comerț al Marii Britanii, relație comercială care depășește mult tranzacțiile cu următorii trei

parteneri comerciali laolaltă (McCrae, 2018). Această generație a crescut crezând că poate la rândul său să se bucure de același lucru — să învețe, să lucreze și să întemeieze familii oriunde în UE. După cum se observă pe website-ul campaniei de rămânere în UE ([www.strongerin.co.uk./](http://www.strongerin.co.uk/)), votul pentru ieșire ar submina toate acestea.¹ Ar lăsa Marea Britanie singură într-o lume conectată și, prin crearea de granițe cu Irlanda, membră UE, ar submina un element central al acordului de pace Good Friday din 1999 cu Irlanda de Nord, acord care a încheiat un conflict lung și sângeros. Și nu doar atât — campania de ieșire din UE nu oferea niciun plan pentru gestionarea aspectelor precum granițele, comerțul și angajamentele deja asumate (Cooper, 2016).

Cu toate acestea, votul „exit“ a câștigat cu 51,9%. Și nu numai atât — cei între 18 și 24 de ani, al căror viitor era în joc, au votat covârșitor ca Marea Britanie să rămână în UE, dar aveau o probabilitate mult mai mică de a vota în comparație cu pensionarii care au venit în forță și au votat covârșitor „exit“ (The Guardian, 2016). Același lucru s-a întâmplat în cazul Irlandei de Nord, unde șase din zece votanți s-au dus la urne (majoritatea celor care au votat au ales să rămână în UE) (BBC, 2016). De ce atât de mulți votanți tineri și atât de mulți nord-irlandezi nu au votat la un referendum care privea propriul lor viitor? De ce ar prefera pensionarii ca Marea Britanie să renege acordul Good Friday și să subestimeze propria siguranță financiară? O posibilitate este că asta s-a întâmplat pentru că oamenii nu au făcut alegerea în funcție de aceste informații așa cum au fost prezentate, ci și-au recadrat alegerea (distorsiunea substituției*****, Kahneman și Frederick, 2002). În loc să se gândească la întrebarea referitoare la modul în care ieșirea din UE ar rezolva problemele societății britanice sau la întrebarea despre cum asta ar oferi alternative la beneficiile includerii în UE, oamenii s-au gândit la altă întrebare. În loc să încerce să sintetizeze informații complexe referitoare la alegerea care ar fi cel mai mult în interesul economiei și siguranței viitorului Marii Britanii, oamenii s-au întrebat care alegere ne reprezintă pe „noi“. Dacă nu s-au putut decide, au rămas acasă (Douthat, 2015; Massie, 2015).

Cum a apărut această recadrare? Acesta este subiectul capitolului de față. Noi sugerăm că oamenii au trecut de la o întrebare la care e dificil de răspuns, care pretinde procesare de informații complexe, la o întrebare cu răspuns simplu, la care se poate răspunde pe baza identității. O abordare bazată pe informații ar implica evaluarea relevanței unor volume mari de date estimate

referitoare la costuri (cât a plătit Marea Britanie către UE), la beneficii (ce a primit Marea Britanie de la UE) și la alternative (schimbările necesare pentru a menține schimburile comerciale, pacea și siguranța granițelor). Pentru asta, oamenii ar trebui, de exemplu, să citească rapoarte pentru a-și da seama dacă apartenența la UE determină creșterea timpului de așteptare pentru serviciile oferite de Serviciul Național de Sănătate (Giuntella, Nicodemo și Vargas-Silva, 2015). Ar trebui să citească rapoarte ca să-și dea seama dacă persoanele nedorite care locuiesc și muncesc în Marea Britanie sunt acolo din cauza UE sau din cauza politicilor prea relaxate pe care Marea Britanie le aplică în ce privește angajarea sau urmărirea oamenilor a căror viză a expirat sau care nu au carte de identitate (Blinder și Markaki, 2018; Goodwin, 2015). Prin opoziție, o abordare bazată pe identitate nu necesită decât ca oamenii să se întrebe ce implică „rămânerea” sau „ieșirea” pentru cine sunt ei și ce ar putea deveni — dacă „să rămână” sau „să iasă” ar fi reprezentativ pentru „noi”.

În acest capitol ne concentrăm asupra încercărilor de persuasiune care îi fac pe oameni să treacă de la raționamente bazate pe informații la raționamente bazate pe identitate. În acest scop, facem diferența între dezinformare și informație. Dezinformarea este conținutul distribuit pentru a determina un anumit raționament sau un curs de acțiune din partea receptorilor, indiferent de veridicitatea sau tendențiozitatea acestui conținut. Prin opoziție, informația este conținutul distribuit pentru informarea receptorilor, iar ceea ce se distribuie se presupune că e adevărat. După cum vom arăta, motivul pentru care dezinformarea are putere de persuasiune și poate chema la arme constă în faptul că se folosește de competența culturală a oamenilor pentru a-i face să treacă de la o procesare bazată pe informații la o procesare bazată pe identitate.

Subliniem pașii pe care îi implică acest lucru — crearea aparenței de întrebare „legitimă” din punct de vedere cultural, cadrarea chestiunii ca fiind determinată de identitate, nu de informații, prezentarea unei alegeri clare, bazate pe identitate și cadrarea alegerilor alternative ca amenințări la adresa identității. La pasul 1, campaniile de dezinformare folosesc competența culturală a oamenilor pentru a recadra subiectele ca întrebări, luând o întrebare care altfel ar fi considerată „ilegitimă” (fiindcă răspunsul e de la sine înțeles) și recadrându-o ca întrebare „legitimă” (la care răspunsul nu e de

la sine înțelese). La pasul 2, campaniile de dezinformare capitalizează competența culturală a oamenilor, astfel încât subiectul e cadrat în termeni fluenți din punct de vedere cultural, folosind imagini, formule și indicatori senzoriali ușor de recunoscut. Odată lucrurile puse în mișcare, campaniile de dezinformare cadrează un curs specific de acțiune ca fiind relevant pentru identitate (ce facem „noi“) și, în plus, sugerează că neîndeplinirea acțiunii relevante pentru identitate reprezintă o amenințare la adresa identității înseși — în cazul ieșirii din UE, rămânerea este cadrată ca pierdere a identității britanice. Astfel, după cum subliniem în acest capitol, lucrurile sunt mai complexe decât simpla lipsă de informații sau prezența informațiilor false. Este vorba despre modul în care mesajele persuasive îi fac pe oameni să se raporteze la propria identitate pentru a înțelege ce implică informațiile în privința acțiunii avute în vedere.

Luăm ca exemplu concret pentru discutarea acestor trei pași Marea Britanie și referendumul din 2016 pentru rămânerea sau părăsirea UE. Referendumul a fost numit Brexit, acronim format din „British“ și „exit“ (ieșire). Această abreviere a contribuit la cadrarea întrebării despre rămânerea sau ieșirea din UE. Denumiri alternative, de exemplu, Brit-in, abreviere formată din cuvintele „Britain“ și „in“ (în) ar fi schimbat cadrarea. În următoarea secțiune operaționalizăm conceptul de competență culturală și explicăm motivul pentru care este relevant în cazul raționamentelor.

Competența culturală, fluența culturală și nefluenața culturală

Dintr-o perspectivă ecologică, traiul în grup este necesar pentru supraviețuire și cultura umană este esențială pentru adaptarea la traiul în grup (Boyd și Richerson, 2005; Cohen, 2001; Haidt et al., 2015; Kurzban și Neuberg, 2005). Traiul în grup implică dezvoltarea „acordajului social“ (sensibilitate la perspectivele celorlalți) și a capacității de „autoreglare“ (capacitatea de control al atenției) (Chiu et al., 2015; Oyserman, 2011, 2017; Shteynberg,

2015). Aceste abilități necesare din punct de vedere cultural sunt baza practicilor culturale care au evoluat pentru a crea soluții „suficient de bune“ pentru problemele de supraviețuire, cum ar fi coordonarea, adecvarea și împărtășirea. Aceste soluții sunt „suficient de bune“, și nu optime, dar, odată dezvoltate, devin „standard“ prin faptul că definesc modul în care „noi“ facem lucrurile — „structurile, practicile, normele și valorile „noastre“ (Cohen, 2001; Oyserman, 2015a). Pătrund în toate aspectele comportamentului, constrâng și permit percepția și raționamentul și oferă o schemă comună de creare de sens în diferite situații (Chiu et al., 2010; Nisbett și Norenzayan, 2002; Oyserman, 2017; Shweder și LeVine, 1984; Triandis, 2007).

Astfel, cultura este parțial un set de rețele asociative de cunoaștere, de coduri tacite de operare, de cadre creatoare de sens prin care oamenii înțeleg lumea, înțeleg ce-și doresc și modul în care pot obține asta. Aceste rețele asociative de cunoaștere, cu rădăcini culturale, oferă modele mentale care le permit oamenilor să dețină competența culturală de a anticipa modul în care e probabil să se desfășoare lucrurile. Fiecare soluție „suficient de bună“ a unei culturi presupune o rețea de cunoaștere care presupune conținut, proceduri și obiective legate de tema ei generală — individualism, colectivism și onoare. Același lucru e valabil pentru fiecare practică a unei culturi. Fiecare practică presupune rețele de cunoaștere, inclusiv cunoaștere specifică, adesea implicită, despre modul în care funcționează lucrurile — cum se îmbracă miresele, ce se mănâncă la micul dejun și așa mai departe. Contextele imediate fac accesibile în fiecare moment un subset al rețelelor culturale de cunoaștere.

Conceptiile determinate cultural modelează procedurile mentale accesibile

Oamenii folosesc subsetul cunoașterii disponibile, determinate cultural, accesibil în momentul judecății, pentru a face predicții automate în privința a

ceea ce urmează să se întâmple. Pentru asta, folosesc proceduri mentale asociate unei concepții culturale accesibile. De exemplu, odată activată o concepție individualistă, oamenii se descurcă mai bine la sarcina de a numi un obiect anume din câmpul vizual (Oyserman, Sorensen, Reber și Chen, 2009, Studiul 3). Acest răspuns sugerează că folosesc o procedură de separare. Prin opoziție, odată activată o concepție colectivistă, oamenii se descurcă mai bine la sarcina de amintire a locului în care s-au aflat obiecte în câmpul vizual (Oyserman et al., 2009, Studiile 1 și 2) și le este dificil să ignore alte informații vizuale (Oyserman et al., 2009, Studiile 4 și 5) sau auditive (Oyserman et al., 2009, Studiile 6 și 7). Acest răspuns sugerează că folosesc o procedură de conectare. Alte studii confirmă aceste predicții bazate pe proceduri ale unei concepții colectiviste accesibile (e.g. Mounsey, Oyserman și Yoon, 2013; Oyserman et al., 2009). După ce se întâmplă asta, oamenii sunt dispuși să plătească mai mult pentru a completa un set (Mounsey et al., 2013 Studiul 1b). Sunt dispuși să accepte opțiuni anterior nedorite dacă un set nu poate fi completat (Mounsey et al., 2013; Studiile 2–4). Găsesc mai dificil să descopere cea mai bună potrivire și să ignore altele, plauzibile, dar nu la fel de bune, în cadrul unei sarcini standardizate de identificare a antonimelor și analogiilor (Oyserman et al., 2009, Studiul 8). Oameni din diferite țări (SUA, Norvegia, Hong Kong, Coreea) și din diferite grupuri etnic-rasiale (e.g. latino, africani-americiani, asiatici sau asiatico-americiani) au adoptat sau nu concepții colectiviste, în funcție de indiciile de moment. În cadrul experimentelor, procedurile mentale pe care le-au folosit oamenii au depins de concepția culturală accesibilă pe moment. Orice determină pe moment ieșirea în evidență a identităților de grup ale oamenilor (e.g. faptul de a fi britanici, de a fi de la țară, de a fi patrioți) ar trebui să declanșeze aplicarea unei concepții colectiviste (Oyserman, 2007). Odată declanșată aceasta, oamenii sunt mai dispuși să se concentreze asupra legăturilor dintre lucruri, ceea ce face mai dificilă identificarea erorilor logice (Oyserman, 2019a).

Cunoașterea culturală modelează ce e fluent și nefluent și, astfel, stilul de raționament

Cunoașterea culturală determină așteptări implicite referitoare la modul în care se vor desfășura lucrurile. E mai simplu pentru oameni să înțeleagă o situație care se desfășoară așa cum se așteptau și e mai dificil să înțeleagă o situație a cărei desfășurare nu respectă așteptările implicite. Un exemplu clasic provine de la Bruner și Postman (1949). Ei au demonstrat că studenții americani recunoșteau mai lent formele (inimă, romb, treflă, pică) când le vedeau în culori nepotrivite (e.g. pică roșie) cunoașterii culturale referitoare la culoarea pe care o au formele din cărțile de joc (ar fi trebuit să fie negre). Studenții din acest studiu au aplicat automat cunoașterea bazată pe cultură. Au făcut asta chiar dacă experimentatorii nu au spus că sarcina era centrată pe cărțile de joc și nu le-au spus că ar trebui să se folosească de cunoașterea pe care o au în privința culorilor formelor din cărțile de joc. Au aplicat automat cunoștințele pe care le aveau. Drept urmare, le-a fost greu să discearnă forma dacă era treflă, dar avea culoarea roșie. Erau predispuși în special să-și aplice cunoștințele bazate pe cultură în cazul primei cărți de joc pe care o vedeau și păreau să continue să facă la fel, în afară de cazul în care mai multe sarcini le demonstau că asta era irelevant. Cultura, desigur, este dinamică, iar acest experiment va avea aceleași rezultate în rândul studenților americani dacă jocul de cărți ar fi la fel de popular pe cât pare să fi fost când acest experiment s-a desfășurat inițial, în anii 1940.

Cunoașterea culturală nu doar face mai simplă procesarea informației fluente din punct de vedere cultural, ci și îi ajută pe oameni să vadă dacă ceva nu e în regulă, determinând trecerea de la procesarea asociativă, intuitivă, la raționamente sistematice, bazate pe reguli, în cazul în care apare ceva neașteptat (Oyserman, 2011). Ca exemplu, să ne gândim la cele patru experimente realizate de Mourey, Lam și Oyserman (2015). În fiecare dintre acestea, un indiciu cultural (variabilă independentă) era prezentat într-o modalitate fluentă din punct de vedere cultural (adecvată așteptării culturale) sau nefluentă din punct de vedere cultural (neadecvată așteptării culturale) într-o situație sau într-un produs. Primul experiment a implicat prezența sau absența unui chenar roz de Valentine's Day sau după Valentine's Day, al doilea și al treilea au implicat evaluarea calității unor fotografii de nuntă, al patrulea a presupus mai întâi alegerea unui format de necrolog pentru o familie și apoi implicarea într-o sarcină cognitivă. Predicția a fost că, pe baza

competenței culturale, procesul de găsire de corespondențe va fi mai simplu decât procesul de identificare a incongruențelor și că această dificultate determinată de competența culturală va provoca o modificare la nivelul raționării sistematice.

Efectul asupra raționării (variabila dependentă) a fost evaluat printr-o sarcină cognitivă creată special pentru a avea și un răspuns intuitiv, și unul bazat pe reguli. Deși răspunsurile bazate pe intuiție nu sunt întotdeauna greșite, acestea induc în eroare în situațiile în care aplicarea procesării bazate pe reguli nu vine ca răspuns firesc, dar este modalitatea corectă de a proceda. În continuare este un exemplu al sarcinii inițiale (preluat din Frederick, 2005): „O undiță de pescuit și momeala costă în total 1,1 dolari. Undița costă cu 1 dolar mai mult decât momeala. Cât costă momeala?” Răspunsul intuitiv, dar greșit, este 0,10 dolari, determinat pe baza informației obținute din scăderea unui dolar din prețul întreg ($1,1 - 1,0 = 0,10$). Răspunsul corect, procesat pe baza regulilor logice, este 0,05 și se obține din informația reflexivă că prețul undiței este cu un dolar mai mare, ceea ce duce la ecuația: $1,1 = n + (n + 1,0)$. Majoritatea oamenilor oferă răspunsul intuitiv 0,10 sau răspunsul reflexiv 0,05, și doar câțiva oferă răspunsuri care nu pot fi codate ca intuitive sau reflexive (alte răspunsuri în afară de 0,10 sau 0,05).

Primul experiment s-a desfășurat în Ann Arbor, Michigan (SUA) și în Hong Kong, China. În fiecare locație, participanții au fost repartizați aleatoriu în patru grupuri — grupurile variaau în funcție de ziua în care au fost abordați (de Valentine's Day sau o săptămână mai târziu) și de culoarea chenarului de pe ecran (roz sau altă culoare). Un grup (cel de fluentă culturală) a realizat sarcina de Valentine's Day pe un ecran cu chenar roz. Celelalte trei grupuri au fost grupuri de control. Au realizat sarcina de Valentine's Day, dar fără chenarul roz. Participanții repartizați aleatoriu în grupul de fluentă culturală au văzut un chenar roz de Valentine's Day (culoarea „potrivită” la momentul „potrivit”). Această corespondență cu așteptarea bazată pe cultură a determinat menținerea procesării intuitive chiar și când era necesară procesarea bazată pe reguli. Într-adevăr, persoanele din grupul de fluentă culturală au avut o probabilitate mai mare de a răspunde greșit decât persoanele din celelalte trei grupuri (între care nu au existat diferențe). Regula nu era dificil de aplicat, era necesar doar ca participanții să vadă că trebuie aplicată. Timpul de care au avut nevoie pentru a răspunde nu a fost

diferit pentru cei care au aplicat o regulă în comparație cu cei care au răspuns intuitiv.

Această descoperire — că până și în situații în care e necesară procesarea sistematică, bazată pe reguli, oamenii rămân la procesarea asociativă, bazată pe intuiție, după ce au fost expuși la indicatori de fluență culturală — a fost reprodusă în trei experimente ulterioare. Este grăitor în aceste experimente faptul că au fost găsite efecte chiar dacă expunerea la indiciul cultural care a declanșat fluența sau nefluența culturală a fost separată de sarcina ulterioară de procesare. În două dintre experimente a fost activată competența culturală referitoare la nunți. În aceste studii, o jumătate dintre participanți au fost repartizați aleatoriu într-un grup care avea sarcina de a evalua opt fotografii de nuntă fluente din punct de vedere cultural și o jumătate au alcătuit grupul care a evaluat opt fotografii de nuntă nefluente din punct de vedere cultural. În fotografiile cu fluență culturală, mireasa era îmbrăcată în alb, mirele, în negru, tortul de nuntă era etajat și glazurat și printre invitați erau domnișoare și cavaleri de onoare. În fotografiile fără fluență culturală, rochia miresei avea verde și mov, și costumul mirelui era cu mov, tortul etajat de nuntă era decorat cu roțițe colorate și nu existau invitați. În ultimul experiment a fost activată competența culturală referitoare la înmormântări și doliu. În acest studiu asupra necrologurilor, o jumătate dintre participanți au fost repartizați aleatoriu în grupul care a evaluat o serie de necrologuri fluente din punct de vedere cultural și cealaltă jumătate a alcătuit grupul care a evaluat o serie de necrologuri nefluente din punct de vedere cultural. În condiția de fluență culturală, participanții au văzut două versiuni ale aceluiași necrolog scris cu un ton trist și care elogia decedatul. În condiția fără fluență, participanții au văzut două versiuni ale aceluiași necrolog scris fără un ton trist și care nu elogia decedatul. Cercetătorii au luat necrologul scris fără tristețe, neelogios, și au creat un necrolog paralel trist și elogios. Astfel, „n-a avut hobby-uri... n-o să i se simtă lipsa“ din original a fost transformat în „a avut numeroase hobby-uri... i se va simți lipsa“. Cercetătorii au realizat două versiuni ale fiecărui necrolog prin rearanjarea ordinii paragrafelor.

În aceste experimente, participanții repartizați aleatoriu în condiția fluentă din punct de vedere cultural au avut o probabilitate mai mare de a aplica procesarea intuitivă decât cei repartizați aleatoriu în condiția fără fluență culturală. După cum demonstrează aceste experimente, experiențele de

fluență și nefluență culturală apar în punctul de legătură între ceea ce competența culturală îi face pe observatori să aștepte (implicit), ceea ce observă aceștia în mod concret și sensul pe care îl extrag din experiențele metacognitive ale simplității sau dificultății (Oyserman, 2011, 2017). Ceea ce determină experiența metacognitivă de simplitate sau dificultate nu este observația în sine, ci potrivirea sau nepotrivirea dintre observație și așteptările bazate pe cultură. Experiența potrivirii sau nepotrivirii implică faptul că persoana are competența culturală necesară pentru a ști (implicit) la ce să se aștepte. Aceste așteptări sunt înrădăcinate în cultura persoanei — ce a învățat în mod explicit sau a asimilat implicit prin observare și practici de socializare. Când mesajele apar în termeni fluenți din punct de vedere cultural, se poate ca oamenii să fie mai susceptibili la dezinformare pur și simplu pentru că mesajul nu declanșează trecerea la procesare sistematică (Oyserman, 2019a).

Competență culturală și mesaje persuasive

Cercetări anterioare asupra fluenței și nefluenței culturale nu au evaluat direct efectele competenței culturale asupra modului în care oamenii procesează mesajele persuasive (pentru o analiză, vezi Oyserman și Yan, 2019). Strategia de procesare pe care o folosesc oamenii este importantă pentru tipul de mesaj pe care îl găsesc persuasiv (Petty și Cacioppo, 1981, 1984). Dacă folosesc un raționament asociativ, există o probabilitate mai mică să observe diferențe la nivelul calității mesajului decât dacă folosesc o abordare sistematică a raționamentului. Mesajele care folosesc o combinație de imagini, formulări, sunete și conținuturi la care oamenii se așteaptă sunt mai fluente din punct de vedere cultural. În această secțiune discutăm modul în care poate fi relevantă fluența culturală a mesajului.

Să ne amintim că e mai puțin probabil ca oamenii să folosească o procesare sistematică în situații fluente din punct de vedere cultural. Noi susținem pe această bază că există o probabilitate mai mică de procesare sistematică în

cazul în care oamenii au de-a face cu tentative de persuasiune fluente din punct de vedere cultural. Întrucât nu raționează sistematic, se poate să nu fie relevantă calitatea în sine a argumentului persuasiv. Ca urmare a lipsei de atenție la calitatea mesajului, se poate ca oamenii să nu facă diferența între informare și dezinformare — între mesaje menite să redea faptele, așa cum sunt cunoscute la momentul respectiv, și informație incorectă factual sau distorsionată. Ei pot să nu facă distincția între mesaje informaționale care au scopul de a constitui fundamentul alegerii și judecății, pe de o parte, și mesaje de dezinformare, care au scopul de a îndrepta spre un anumit curs de acțiune, pe de altă parte. Pentru ca un mesaj să fie fluent din punct de vedere cultural, este necesar ca acesta să conțină o combinație de imagini, formulări și conținuturi pe care oamenii se așteaptă tacit să le vadă sau să le audă într-o anumită situație.

După cum vom detalia în continuare, fluența culturală activează procesarea bazată pe identitate, nu pe informații. Când oamenii folosesc procesarea bazată pe informații au acces și la strategii asociative, și la strategii sistematice. Prin opoziție, când folosesc procesarea bazată pe identitate, accesul la strategii sistematice este efectiv blocat. Pentru a explica de ce se întâmplă asta, noi aplicăm logica comunicării, expusă în continuare.

Logica comunicării

Conform convențiilor de utilizare a limbajului, oamenii presupun în general că persoanele care expun un mesaj au scopul de a informa (maximele comunicării ale lui Grice sau „logica comunicării“, Schwarz, 2014). Cu alte cuvinte, oamenii presupun că cei care expun un mesaj împărtășesc un conținut pe care îl consideră adevărat, nepărtinitor și potențial util pentru a sta la baza judecăților și luării de decizii, chiar dacă uneori aceste persoane înțeleg greșit mesajul și, fără să vrea, dezinformează — adică distribuie conținut neadevărat sau părtinitor. Conform acestor convenții ale utilizării limbajului, în afară de cazul în care au motive să fie suspicioși, oamenii

pornesc de la premisa că emițătorii încearcă să informeze — să spună clar ceva relevant și ce audiența lor încă nu știe (Schwarz, 2014).

Logica comunicării și intenția de comunicare

Logica comunicării servește oamenilor când emițătorul și receptorul împărtășesc obiectivul de a informa. Dat fiind că raționamentul lor este modelat de logica comunicării, oamenii fac presupuneri (adesea implicite) referitoare la informație pe baza modului în care este comunicată (Gilbert, 1991; Gilbert, Krull și Malone, 1990; Schwarz, Strack, Hilton și Naderer, 1991; Schwarz, 2014; Schwarz și Sudman, 2012; Sudman, Bradburn și Schwarz, 1996). Procedează astfel indiferent dacă emițătorul a avut intenția de a-i face să ajungă la aceste inferențe și adesea fără să fie conștienți de sursa inferenței (Schwarz et al., 1991).

Însă logica comunicării poate și să ascundă intențiile emițătorilor care nu au scopul de a oferi informații în vederea judecății și alegerii, ci au scopul de a modela judecata și de a determina un anumit rezultat (o judecată, o alegere). Deși veridicitatea mesajului și părtinirea sunt relevante când emițătorii au obiectivul de a informa, sunt irelevante când emițătorii au scopul de a modela judecata și de a determina o anumită alegere. Așa cum am observat în primele paragrafe, folosim termenul „dezinformare” pentru a descrie această formă din urmă a conținutului mesajului distribuit de emițători care nu au intenția de a informa, ci de a modela judecata și decizia receptorului, indiferent de concludența mesajului pe care îl transmit. Veridicitatea și părtinirea sunt irelevante în cazul mesajelor de dezinformare, întrucât nu contează dacă conținutul este adevărat sau nepărtinitor; tot ce contează este producerea răspunsului dorit (Weedon, Nuland și Stamos, 2017). Noi înaintăm ideea că oamenii au o probabilitate foarte mică de a observa dezinformarea dacă apare într-un mod fluent din punct de vedere cultural pentru că, în aceste situații, e mai puțin probabil să fie suspicioși, să observe că ceva nu e în regulă și să

treacă la procesarea sistematică.

Logica comunicării și întrebările „legitime“ (versus cele „nelegitime“)

Dat fiind că oamenii fac presupuneri pe baza modului în care informația este comunicată, emițătorii pot să provoace îndoieli doar prin punerea de întrebări. Conform logicii comunicării, recipientii întrebărilor în general presupun o intenție de cooperare. În ce privește întrebările, intenția de cooperare implică faptul că persoana care comunică pune o întrebare pentru că e posibil să existe mai mult de o opțiune. Existența mai multor răspunsuri posibile este ceea ce face ca o întrebare să fie legitimă. Dacă nu există decât un răspuns posibil, întrebarea nu e legitimă. Dar posibilitatea ca întrebarea să nu fie legitimă este în general trecută cu vederea când oamenii presupun intenția de cooperare. Acesta este motivul pentru care provoacă confuzie întrebarea: „Cine e îngropat în mormântul lui Grant?“ (întrebare care include răspunsul, Grant); să fie o întrebare legitimă ar însemna să nu fie posibil doar un răspuns. Ceea ce implică faptul că Grant nu e îngropat în mormântul lui Grant, că singurul răspuns posibil (generalul Grant) nu e singurul răspuns posibil. Dacă obiectivul nu este de a informa, ci de a dezinforma — de a schimba judecata, nu de a oferi informații pe care să se bazeze, atunci punerea unei întrebări poate fi un prim pas pentru schimbarea judecății. Când sunt întrebați „Cine e îngropat în mormântul lui Grant?“, oamenii răspund adesea „Nu știu, cine?“, excluzând că Grant e înmormântat acolo. Dar, desigur, aceasta nu e o întrebare legitimă — Grant este înmormântat în mormântul lui Grant. Noi propunem ideea că astfel de întrebări ilegitime, care nu se bazează pe absența unui unic răspuns, au intenția de a semăna îndoială și de a-i face pe oameni să fie deschiși în fața oricărei alternative posibile.

În 2016, întrebarea: „Ar trebui Marea Britanie să iasă din UE?“ nu a fost, cel puțin inițial, o întrebare în mod clar legitimă. La urma urmei, dacă părăsirea

UE ar putea fi considerată o întrebare legitimă, asta ar însemna că s-ar putea să nu fie adevărate lucrurile pe care le credeam adevărate, așa că poate apartenența la UE e o problemă, altfel de ce să fie pusă întrebarea? Prin faptul că a pus întrebarea pe buletinul de vot, campania de secesiune a reușit să facă întrebarea să pară întrucâtva legitimă. Dincolo de punerea întrebării pe buletinul de vot, campania de ieșire din UE ar fi putut folosi mesaje de informare sau de dezinformare, după cum, desigur, ar fi putut face și campania de rămânere în UE. În următoarea secțiune subliniem ce anume este motivația bazată pe identitate. Apoi vom folosi motivația bazată pe identitate pentru a explica atractivitatea campaniilor de dezinformare, utilizând exemple din campania Brexit (am căutat, dar nu am găsit exemple în campania de rămânere în UE).

Motivația bazată pe identitate

Construirea dinamică, interpretarea experienței și disponibilitatea pentru acțiune

Teoria motivației bazate pe identitate este o teorie cognitivă consacrată a autoreglării, care prezice că oamenii preferă să confere sens situațiilor și să se comporte în moduri congruente cu identitatea lor socială și personală importantă (Oyserman, 2007, 2009, 2015b). Identitatea socială este identitatea care are legătură cu apartenența la grup — faptul de a fi patriot, naționalist, britanic, londonez, european, bărbat, părinte, plătitor de taxe. Aceste identități pot avea legătură cu diverse conținuturi și, astfel, se pot suprapune identităților personale — persoană responsabilă fiscal, mândră și loială. Identitățile sociale pot avea ca referent și conținut semantic (ce

valorizăm, credințele noastre), și conținut senzorial — cum arătăm, cum sună vocea noastră, care sunt gusturile care ne plac. Oamenii au multe identități sociale și personale în trecut, în prezent și poate în viitor, disponibile în memorie; aceste identități nu au o structură organizațională predefinită (Oyserman, Elmore și Smith, 2012), ci oamenii sunt influențați de acele identități care sunt accesibile („în minte“) în momentul în care fac o judecată, în cazul în care aceste identități par relevante pentru sarcina pe care o au de îndeplinit.

Construire dinamică

Pentru a-l parafraza pe William James (1890), gândirea (despre propria persoană) are scopul acțiunii. Dat fiind că acțiunea necesită evaluarea posibilităților și constrângerilor prezente într-o anumită situație, identitățile disponibile în minte și ceea ce aceste identități par să însemne se află în rezonanță cu aspectele semnificative pe moment și în general ale contextului (pentru o analiză, vezi Oyserman et al., 2012). Oamenii nu extrag pur și simplu din memorie ce înseamnă o identitate, ci și fac inferențe pe baza contextului în ce privește ce anume trebuie să însemne o identitate în condițiile aspectelor situației de față. În acest sens, identitățile sunt construite dinamic în prezent — o identitate aparent constantă poate implica acțiuni diferite în contexte diferite. Astfel, în prezent, faptul de a fi britanic poate fi parte a faptului de a fi european — când călătorim fără să fie nevoie de vize, dar ar putea și să fie opus faptului de a fi european — când oamenii din alte țări își înscriu copiii la școala din zonă. Sensibilitatea oamenilor la implicațiile situației imediate este un aspect esențial, nu o eroare de design. Sensibilitatea la contextul social le permite oamenilor să facă inferențe în privința a ceea ce e probabil să facă alți oameni la fel ca ei, ce strategii funcționează pentru ei și ce inferențe pot fi făcute când lucrurile merg bine și când întâmpină dificultăți (Oyserman et al., 2017).

Aplicând logica pe care am subliniat-o în secțiunea despre evoluția culturii,

mesajele distribuite în cadrul grupului ar trebui să pară mai credibile — membrii grupului au aceleași valori și au o probabilitate mai mică de a aduce vătămări în comparație cu cei care nu fac parte din grup. Dintr-o perspectivă evoluționistă, capacitatea de a recunoaște cine face parte din grup este esențială. Interacțiunea cu cei dinăuntru grupului e sigură, ei pot fi abordați. E improbabil ca membrii grupului să înșele sau să reprezinte o amenințare, ceea ce reduce necesitatea temerii, suspiciunii și precauției (Brewer, 1979; Platow, Foddy, Yamagishi, Lim și Chow, 2012). Cu cei din același grup persoana poate să-și lase garda jos, dar cum își poate da seama cine distribuie mesajul? Simțurile pot oferi indicii — oamenii ca „mine“ vorbesc într-un anumit fel, folosesc anumite formule, au anumite accente și oamenii ca mine „arată“ într-un anumit mod, se îmbracă într-un anumit stil, se regăsesc în anumite imagini, oamenii ca „mine“ au aceleași gusturi, valori și dorințe. Astfel, mesajele din cadrul grupului au o probabilitate mai mare de „a părea adevărate“ și de „a suna bine“; e mai probabil să pară familiare și să fie fluente din punct de vedere cultural.

Disponibilitate procedurală

Crearea de sens sau „disponibilitatea procedurală“ este disponibilitatea de a înțelege noile situații în moduri bazate pe identitățile activate. De exemplu, când sunt activate identitățile sociale bazate pe un „noi“ colectiv, există o probabilitate mai mare ca oamenii să folosească proceduri mentale de conectare și relaționale, iar dacă sunt activate identitățile personale de tip „eu“, oamenii au o probabilitate mai mare de a folosi proceduri mentale de separare și diferențiere (Oyserman, 2007). Acest lucru implică faptul că, în mesajele transmise de campaniile de dezinformare al căror conținut include identități sociale, avem de-a face cu declanșatorul unui tip de gândire bazat pe conexiuni și asociații, și nu pe concentrarea asupra unui punct esențial. Prezentăm exemple ale acestui proces în secțiunea Campanii de dezinformare și motivații bazate pe identitate.

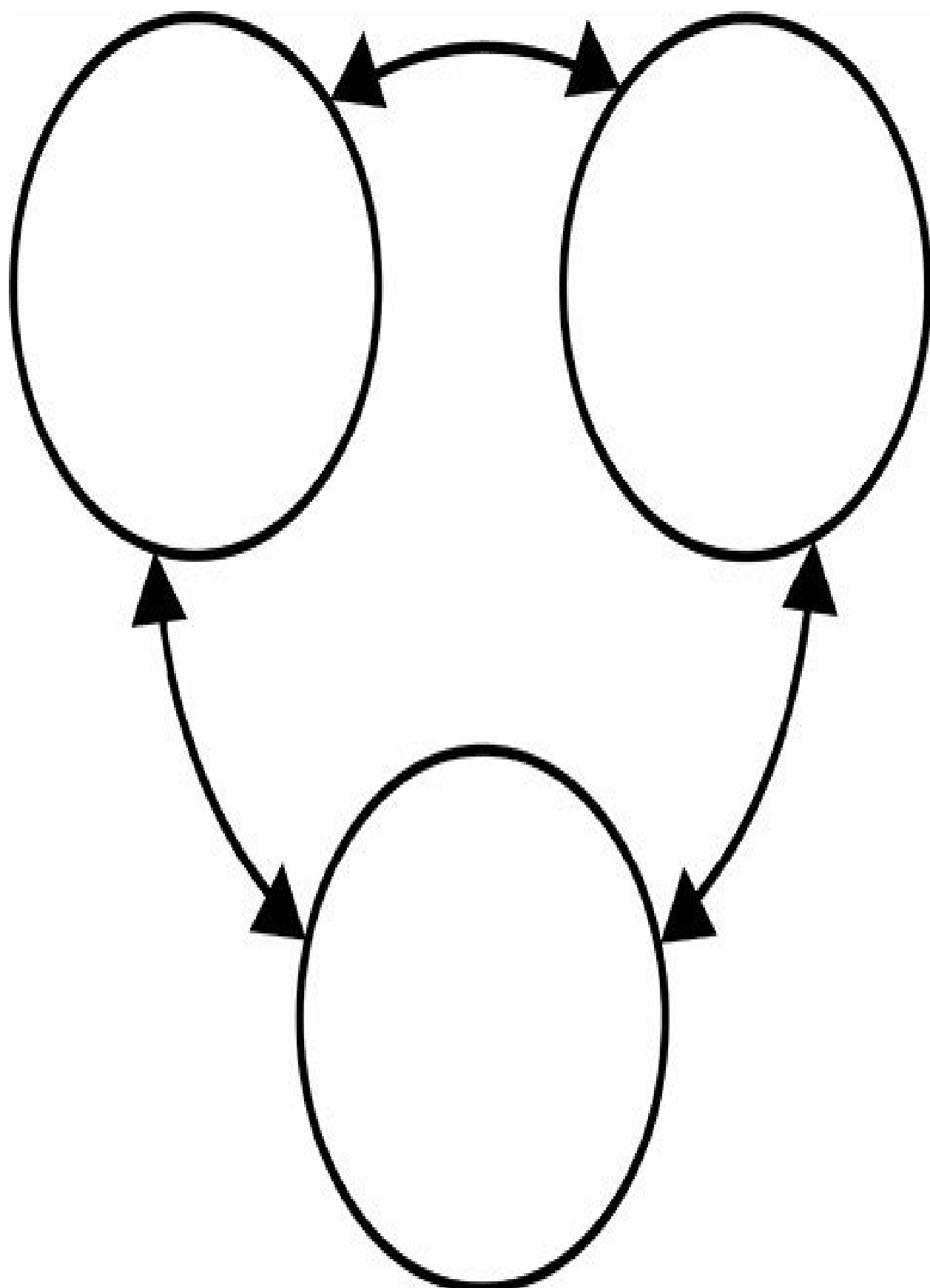
Disponibilitatea de acțiune

Disponibilitatea de a acționa în situații noi sau ambigue în modalități congruente cu propria identitate este numită „disponibilitate de acțiune”. Dacă o anumită acțiune este congruentă cu identitatea „mea” sau „a noastră”, persoana va persevera dacă apar dificultăți la început sau pe parcurs. Invers, dacă o anumită acțiune este irelevantă pentru identitate sau chiar incongruentă cu aceasta, este mai probabil ca oamenii să interpreteze diferit dificultățile de a începe sau de a continua o anumită acțiune. În aceste cazuri, dificultatea înseamnă că acțiunea oricum nu este pentru „mine” (Elmore și Oyserman, 2012, Oyserman, 2019b). Cercetările anterioare s-au concentrat asupra interacțiunii dintre identitățile sociale și implicarea în activități școlare — studiu, participare la discuția de la ore, concentrarea atenției, solicitarea ajutorului, împrumutarea de cărți de la bibliotecă (e.g. pentru analize, vezi Oyserman et al., 2012; Oyserman, 2019b). Deși fără a fi fost evaluată direct, implicația este că acțiunile congruente cu identitatea reprezintă confirmări ale identității, iar lipsa de acțiune este o amenințare la adresa identității. De exemplu, dacă elevii buni studiază, să studiezi înseamnă că ești un elev bun; dacă nu înveți înseamnă că ești altceva. În cazul campaniei Brexit, identitățile sociale, inclusiv faptul de a fi britanic și ecologist, au fost puse în legătură cu acțiunea specifică de a vota „ieșirea”. Implicația este că, dacă asta facem „noi”, cineva care nu votează „ieșirea” s-ar putea să nu facă cu adevărat parte din grup și, mai rău, s-ar putea să contribuie la dispariția grupului și a ceea ce reprezintă acesta.

După cum se ilustrează în Figura 10.1, fiecare dintre cele trei componente bazate pe identitate ale motivației (construirea dinamică a identității, disponibilitatea de a acționa și crearea de sens în modalități congruente cu identitatea) operează în tandem. Asta înseamnă că indiciile de trecere la acțiune nu doar că declanșează acțiunea, ci ajung să fie semnificative și pentru crearea de sens și identitate. Astfel, într-un anumit context, actualizarea unei identități implică și activarea implicațiilor acesteia pentru

crearea de sens și pentru trecerea la acțiune. Același lucru e valabil dacă e actualizată disponibilitatea de a trece la acțiune (sunt activate implicațiile acțiunilor pentru identitate și pentru crearea de sens) sau dacă e actualizată o modalitate de creare de sens unei experiențe (sunt activate implicațiile creării de sens pentru identitate și pentru acțiune).

Figura 10.1. Trei componente interdependente



Construcție dinamică

Cine suntem „noi“?

Disponibilitate de trecere la acțiune

Ce facem „noi“?

Disponibilitate procedurală

Cum gândim „noi“?

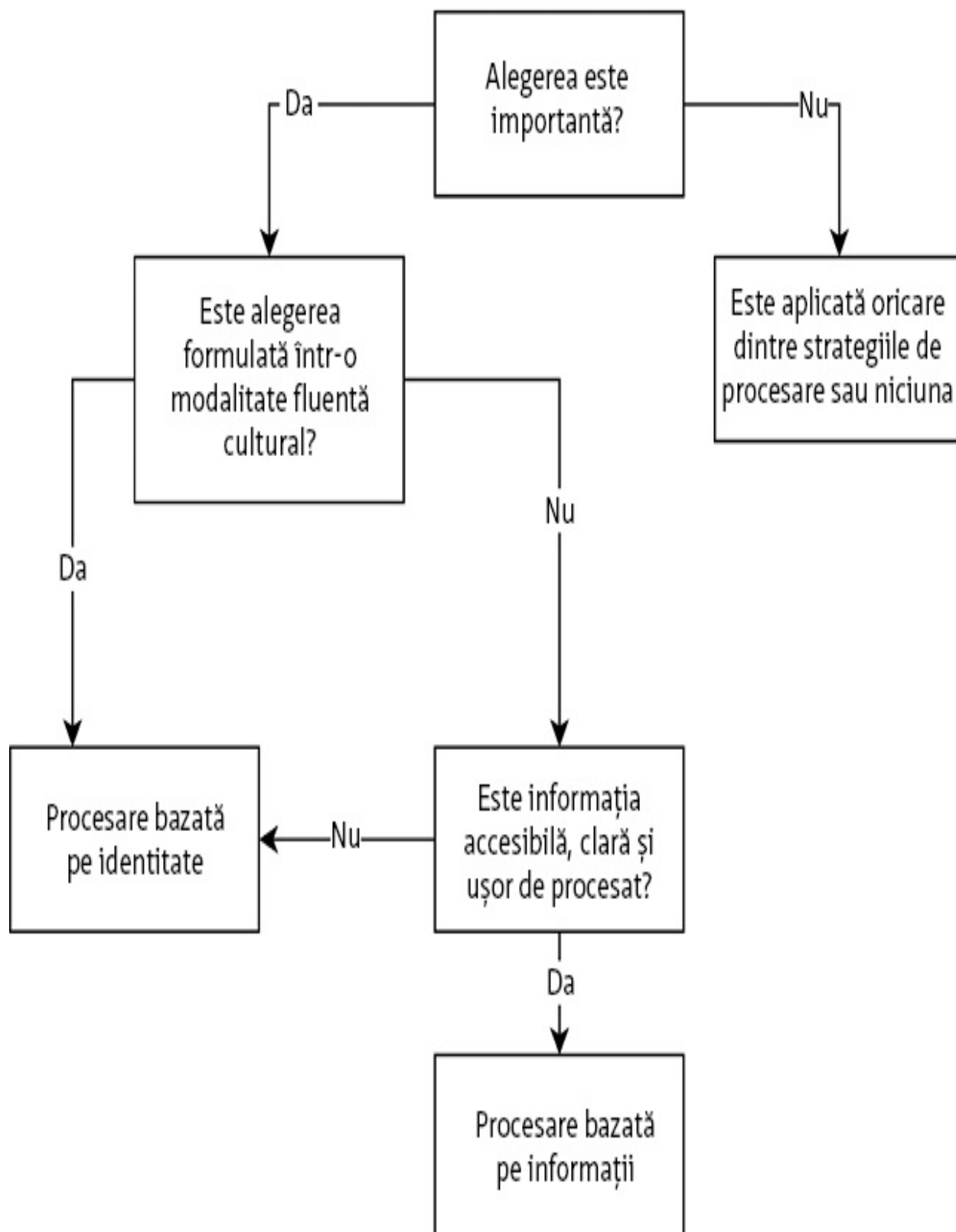
Stabilitatea identității este o ficțiune utilă

Deși construcția dinamică este un atribut esențial al funcționării motivației bazate pe identitate, se poate ca oamenii să nu aibă experiența propriilor procese identitare sau motivaționale în acest mod flexibil. În general, oamenii au experiența propriei identități ca fiind stabilă în timp și spațiu (Oyserman, 2019b). Această convingere este utilă din mai multe motive. În primul rând, le permite oamenilor să facă predicții despre preferințele lor viitoare pe baza a ceea ce preferă acum prin punerea semnului egal între „eu“ din prezent și „eu“ din viitor, considerându-le ca fiind în esență același „eu“ (Oyserman, 2019b). În al doilea rând, asta facilitează alegerea între mai multe acțiuni disponibile, adică îndeplinirea unei acțiuni în prezent spre beneficiul unui „eu“ viitor — să-și facă temele (Nurra și Oyserman, 2018) sau să economisească pentru ieșirea la pensie (Lewis și Oyserman, 2015) — ceea ce are sens dacă „eu“ din prezent și „eu“ din viitor sunt identice. În al treilea rând, prin amplificarea nivelului de certitudine, minimizează gradul în care oamenii tind să caute informații care să le confirme alegerile și modalitatea de creare de sens care stau la baza identității.

Important, dar dificil: trecerea de la procesarea bazată pe informații la procesarea bazată pe identitate

Desigur, nu e obligatoriu ca oamenii să folosească procesarea bazată pe identitate; pot apela (și chiar o fac) la procesarea bazată pe informații. Procesarea bazată pe informații presupune utilizarea informațiilor disponibile ca ghid pentru judecată și alegere rațională. Este probabil ca oamenii să folosească procesarea bazată pe informații atunci când informațiile sunt ușor de accesat, clare și limitate; în aceste cazuri, procesarea computațională (e.g. calcularea riscurilor și recompenselor) este posibilă (Reyna, 2004). Procesarea bazată pe informații poate fi rapidă. De exemplu, evaluarea drumului celui mai scurt se poate face printr-o aplicație de supraveghere online a traficului; costul mai mare poate fi găsit prin compararea prețurilor. Însă, adesea, informațiile utile pentru alegeri și pentru judecăți raționale nu sunt astfel, în special când alegerea este importantă pe termen lung, dar rezultatele pe termen lung sunt complexe, incerte și dificil de procesat. În aceste cazuri, încercarea de a aplica o regulă computațională pentru procesarea informației nu numai că este dificilă, dar s-ar putea să nu fie posibilă, și astfel e necesar să se folosească o altă strategie. Noi sugerăm că prezentarea informației într-o formă fluentă din punct de vedere cultural facilitează trecerea de la o întrebare bazată pe informații, pentru care e dificilă găsirea unui răspuns, la o întrebare bazată pe identitate, la care răspunsul este simplu. Rezumăm acest proces în Figura 10.2.

Figura 10.2. Fluența culturală ca declanșator al procesării bazate pe identitate



Social media și diseminarea dezinformării

Platformele de social media au scopul de a strânge oamenii laolaltă și de a împărtăși conținuturi relevante pentru identitate. Aceste platforme pot părea gratuite și prietenoase — oamenii au senzația că ei înșiși aleg în ce să se implice, că alegerile lor nu sunt constrânse și că pot să aleagă cu cine se afiliază. Însă, prin implicarea voluntară într-un cadru de grup care pare prietenos și sigur, oamenii oferă totodată organizatorilor platformei o bază amplă de informații despre ei înșiși și despre rețelele lor. Ei și informațiile lor devin un produs pe care organizatorii platformei îl pot vinde. Această combinație — flux de informații printr-o legătură cu un „prieten“ personal sau un „follower“ și existența unor informații ample — fac platformele sociale ideale pentru campaniile care încearcă să răspândească dezinformare fluentă cultural ținută spre identități sociale importante. Intenționat sau nu, utilizatorii își oferă către corporații datele și alți indicatori comportamentali foarte semnificativi, în schimbul capacității de a se conecta și de a împărtăși informații (Redazione, 2018). O mare parte dintre aceste date pot fi transformate cu ușurință în indicatori demografici pentru publicitate. Facebook, de exemplu, generează o mare parte din venituri din publicitate, obținând în 2018 peste 55 de miliarde de dolari (Facebook, 2019). Facebook a recunoscut că i-a permis companiei Cambridge Analytica să preia informațiile a aproximativ 87 de milioane de utilizatori, inclusiv rețelele lor sociale (Kang și Frenkel, 2018).

La bază, platformele social media sunt rețele foarte eficiente de publicitate. Algoritmurile lor au scopul de a amplifica interacțiunea cu conținutul și timpul petrecut pe platformă prin direcționarea atenției spre conținut stimulator, personalizat, creat pe potrivea perspectivei utilizatorului sau spre conținut bazat pe emoții (Barberá, Jost, Nagler, Tucker și Bonneau, 2015; Kramer, Guillory și Hancock, 2014). Facebook creează profiluri detaliate ale utilizatorilor, inclusiv date de recunoaștere facială, locația geografică, interese, încadrare demografică, comportamente și hărți ale rețelelor sociale;

prin faptul că permite interacțiunea, poate recolta informațiile cu care contribuie oamenii, conținuturile specifice cu care utilizatorii interacționează și ce fac cu aceste conținuturi (Facebook, fără dată). Toate aceste date pot fi folosite pentru conceperea și diseminarea eficientă a dezinformării (Facebook, 2019; Shochat, Shniberg, Hirsch și Tagiman, 2009). Tehnicile big data permit extragerea de măsurători specifice — demografice și psihologice, cum ar fi trăsături de personalitate și alte caracteristici — din aceste informații (Kosinski, Stillwell și Graepel, 2013).

Motivația bazată pe identitate și campaniile de dezinformare

Ilustrăm modul în care funcționează acest proces de dezinformare revenind la exemplul campaniei Brexit. Pentru a reuși, campania de secesiune trebuia să îndeplinească două lucruri: să convingă o parte dintre votanți să voteze „ieșire“ și să convingă alți votanți să stea acasă și să nu voteze deloc. Pentru a face asta, campania Brexit a folosit diseminarea țintită a dezinformării (conținut fals sau manipulat creat nu pentru a informa, ci pentru a determina o anumită acțiune). Campania a redus probabilitatea ca votanții să observe că mesajele erau de dezinformare, prin utilizarea de materiale fluente din punct de vedere cultural (reducerea probabilității de a utiliza procesarea sistematică) și prin apelul la identități sociale (creșterea probabilității de aplicare a procedurilor mentale de raportare la colectivitate, adică procesarea prin conexiuni și asociații). Campania a crescut probabilitatea ca mesajele de dezinformare să fie acceptate, cadrând judecata și alegerea în termeni de identități sociale (cum gândim „noi“, alegerile pe care le facem „noi“), determinând probabil și disponibilitatea la acțiune, și un cadru colectivist (care ar trebui să crească receptivitatea la intenția de comunicare a emițătorului, e.g. Haberstroh, Oyserman, Schwarz, Kühnen și Ji, 2002). Apoi, campania Brexit a cadrat o anumită acțiune (votarea „ieșirii“) ca fiind relevantă pentru identitate în cazul unora dintre votanți. În cazul altora,

campania s-a concentrat asupra subminării încrederii în identitatea actualizată sau în privința lucrurilor pe care aceasta le-ar implica pentru comportament. Această reformulare pe baza fluentei culturale a avut succes sub două aspecte. În primul rând, a făcut ca Brexit să fie o întrebare legitimă. În al doilea rând, i-a scutit pe oameni de la a trebui să analizeze estimări complexe, contradictorii și incerte ale costurilor financiare inerente rămânerii sau ieșirii și le-a permis ca, în loc de asta, să se întrebe ce înseamnă votul „rămânere“ sau „ieșire“ (sau chiar participarea la vot) pentru cine sunt ei. Desigur, această reformulare care trece de la alegerea bazată pe informații la alegerea bazată pe identitate nu ar fi putut funcționa dacă oamenii n-ar fi avut deja tendința de a face alegeri congruente cu propria identitate, de a trece la acțiuni congruente identității și de a da sens experiențelor lor în modalități congruente identității.

Pentru a obține aceste rezultate, două campanii diferite de secesiune, „BeLeave“ și „Vote Leave“, au angajat o companie specializată în relații digitale de afaceri care s-a ocupat de mesajele de persuasiune transmise prin platformele media (House of Commons, 2019). Această companie, Aggregate IQ (AIQ), este o firmă nord-americană ai cărei fondatori sunt specializați în puterea de persuasiune pe care o au platformele digitale, inclusiv Facebook; reprezenta unul dintre principalii actori din domeniul persuasiunii în masă prin media digitale, mai ales în sfera politică. AIQ a dezvoltat produse software pentru grupul SCL, o mare „agenție globală de management electoral“, cunoscut mai ales pentru subsidiara Cambridge Analytica. Instrumentele dezvoltate de AIQ au fost folosite și în alegerile din America de Nord de către grupul SCL, care a lucrat pentru Partidul Republican (House of Commons, 2018).

Pentru a convinge votanții britanici la referendumul Brexit, AIQ și-a folosit cunoașterea despre modul în care operează Facebook pentru a genera mii de conținuturi diferite pe această platformă (Facebook, fără dată). Compania a aplicat o strategie de dezinformare bazată pe modelarea mesajului în funcție de identitatea receptorilor. Arhitectul-șef al acestei operațiuni a formulat astfel teza subiacentă: „Noi folosim date pentru a construi o imagine a identității persoanei... creăm intervenții personalizate — intervenții informaționale sau comunicaționale — care îi vor schimba comportamentul într-un mod destinat acelei persoane“ (Cadwalladr, 2018; Redazione, 2018).

AIQ s-a consultat cu campaniile în privința eficacității conținutului preexistent și a creat independent volume mari de conținuturi noi. A oferit aceste conținuturi prin echipe interne de diseminare, folosindu-se de modalități sofisticate de cadrare demografică și creare de profiluri — adică știau exact cui se adresează fiecare conținut în parte și fiecare individ primea mesaje alese anume în funcție de propriile tendințe personale (House of Commons, 2018). În acest mod, cetățenilor britanici li s-a oferit conținut care părea „bun“, „sună adevărat“ sau „le vorbea“ într-o manieră clară și firească. Dar ceea ce utilizatorilor neinformați de social media le părea a nu fi nimic altceva decât imagini și sloganuri care atrag atenția erau de fapt cadre culturale fluente, atent construite pentru a transmite o chemare la acțiune modelată în funcție de identitatea receptorilor.

Efectele cadrării motivaționale fluente cultural, bazate pe identitate: campania Brexit

Folosirea avantajului oferit de motivația fluentă cultural, bazată pe identitate, pentru a crește numărul voturilor în favoarea „ieșirii“

Figura 10.3 (ceașca de ceai) și Figura 10.4 (ursul polar) sunt două exemple care ilustrează conținuturile menite a stimula votul pro-„ieșire“. Mesajul din imaginea cu ceașca de ceai preia o identitate „britanică“ și construiește dinamic un anumit conținut pe baza acesteia. Folosește imagini fluente din

punct de vedere cultural (Big Ben, cabina roșie de telefon) și indicatori senzoriali (gustul implicit al ceaiului britanic și sonoritatea „cuppa“), semnificative pentru faptul de a fi „britanic“ și creează un sens nou, conform căruia cel mai bun lucru pe care îl poți face ca „britanic“ este să votezi ieșirea din UE. Acest mesaj le este adresat oamenilor pentru care faptul de a fi britanici este multifățetat și trezește cu ușurință oarecare nostalgie. Adică nu doar că se folosește cuvântul „cuppa“, care are sensul informal de „ceașcă de ceai“, indicator lingvistic în discursul cotidian al englezilor, ci sintagma „not my cuppa [tea]“, care înseamnă „nu e pentru mine“ adaugă și mai multă fluență culturală. Nu doar că mesajul ceștii de ceai cadrează un anumit curs de acțiune pentru receptori, ci este și atât de nostalgic, încât e improbabil să fie considerat relevant de alte audiențe. Astfel, e improbabil să mobilizeze audiența care ar putea vota „rămânerea“.



Figura 10.4. Cadru ecologist fluent cultural, conform căruia ieșirea din UE este necesară pentru menținerea unei identități ecologice

Sursă: retipărire preluată din House of Commons sub Open Parliament License v3.0

Mesajul ursului polar cadrează o altă identitate, ecologistă, și sugerează că aceasta implică o acțiune specifică — ieșirea din UE. Este adresată oamenilor pentru care sunt mai la îndemână alte identități sociale decât cele ale englezilor nostalgici. La fel ca în cazul mesajului ceștii de ceai, oamenii care nu tind să se considere ecologiști au o probabilitate redusă de a procesa informația ca fiind relevantă — în cel mai bun caz, le pare naivă. Neavând cadrul care să confere sens ursului polar, este improbabil să răspundă oricum la acest mesaj.

THE EUROPEAN UNION
WANTS TO KILL OUR CUPPA

LEARN MORE

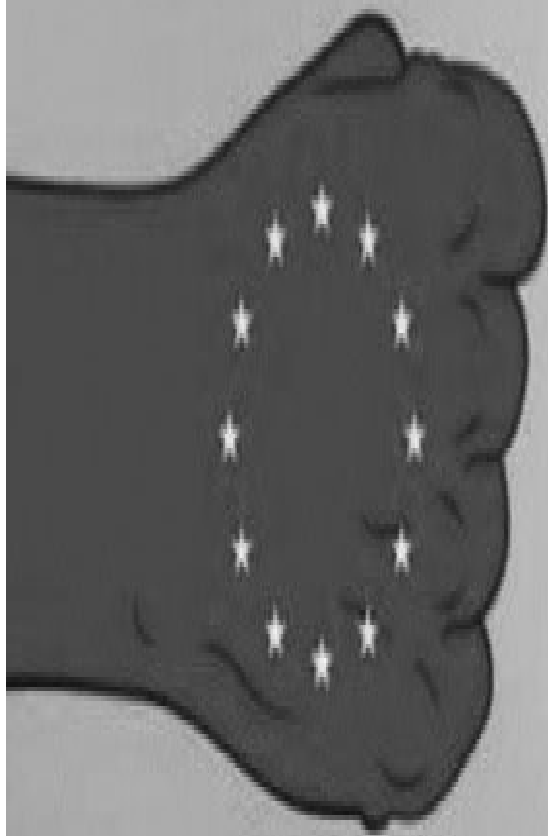


Figura 10.3. Cadru identitar fluent cultural al identității britanice nostalgice, conform căruia părăsirea UE este necesară pentru menținerea identității britanice

Sursă: retipărire preluată din House of Commons sub Open Parliament License v3.0

Folosirea avantajului oferit de motivația fluentă cultural, bazată pe identitate, pentru a submina votul pentru rămânerea în UE

Pentru a convinge potențialii votanți pro-„rămânere“ să stea acasă, campania Brexit a avut două opțiuni. Putea să amplifice îndoiala că votarea „ieșirii“ ar fi fost de fapt un lucru pe care îl facem „noi“ sau putea amplifica îndoiala că votul în sine ar fi fost ceva ce facem „noi“. Figura 10.5 (călătoria cu avionul) oferă un exemplu al tipului de conținut menit a submina încrederea că votul ar fi fost un lucru pe care îl fac „eu“ sau pe care îl facem „noi“. Acest mesaj cadrează două identități posibile — o identitate britanică rațională, frugală, și o identitate ecologistă. Cadrul identitar britanic este în anumite privințe similar cadrului nostalgic al ceștii de ceai; în acest caz reamintește anii de austeritate de după război. Precum în cazul cadrului ceștii de ceai, nu există nicio ambiguitate în ce privește chemarea la acțiune — votează „ieșire“. Spre deosebire de acesta, cadrul identității ecologiste pune întrebarea dacă rămânerea sau ieșirea este cea mai bună alegere pentru mediu. La fel ca multe alegeri ecologiste — este mai bine să spălăm plasticul reciclabil (risipind apă) sau să-l aruncăm într-o groapă de gunoi (risipind energie)? Sau e mai simplu să nu alegem? Pentru ecologist, mesajului călătoriei cu avionul subminează convingerea că votarea „rămânerii“ sau „ieșirii“ este acțiunea congruentă din

punct de vedere identitar. Mai mult, cu aluzia de posibilă corupție, acest mesaj subminează convingerea că liderii politici au vreun alt interes în afară de cel propriu.



Figura 10.5. Cadru fluent cultural care echivalează susținerea UE cu susținerea corupției, ceea ce subminează încrederea că identitatea europeană și identitatea britanică ar fi compatibile

Sursă: retipărire preluată din House of Commons sub Open Parliament License v3.0

Lipsește claritatea în privința acțiunii care trebuie îndeplinită și mesajul reduce probabilitatea oricărei acțiuni. Expunerea la acest mesaj reduce, de asemenea, probabilitatea ca ecologiștii să considere faptul de a vota ca fiind congruent cu propria identitate, și în același timp amplifică probabilitatea ca pensionarii nostalgici să vadă lucrurile așa. Subminează convingerea că votul în sine ar fi congruent cu identitatea, dat fiind că mesajele dintr-o sursă coruptă au o probabilitate foarte mică de a oferi informații utile pentru ceea ce oamenii ca „mine“ ar trebui să facă.

După vot: efecte pe termen lung

Dat fiind că declanșează procesarea bazată pe identitate, fluentă din punct de vedere cultural, efectele campaniilor de dezinformare sunt probabil durabile. Interacționând cu indicatori cu fluentă culturală a identității sociale, este probabil ca oamenii să-și asume o identitate. Această identitate declanșează imediat o acțiune și ajunge să constituie cadrul pentru judecăți ulterioare. În cazul campaniei Brexit, acțiunea imediată este schimbarea comportamentului de vot. Faptul de a fi britanic nu este neapărat antitetic faptului de a fi european, dar odată lucrurile cadrate astfel, este probabil ca oamenii să facă această asociație ori de câte ori sunt actualizați indicatori ai identității britanice. Prin crearea unei legături între o acțiune și procesarea bazată pe identitate, dezinformarea declanșează și menține un stil asociativ de procesare. Întrucât identitățile sociale declanșează procedura de conectare și relaționare colectivă, oamenii au senzația că acțiunea sugerată prin manipulare este alegerea relevantă pentru propria identitate. Devine modul în care „noi“ acționăm, fiind implicit faptul că se potrivește valorilor „noastre“. După cum am descris în secțiunea despre construirea dinamică, dezinformarea nu are nevoie să se bazeze pe asociații deja disponibile între identitate și acțiune. Aceste asociații pot fi construite în context. Cu toate acestea, după ce au fost construite și utilizatorii au interacționat cu ele, ori de

câte ori acea identitate este actualizată, acțiunile asociate și valorile implicite vor fi actualizate la rândul lor (după cum se ilustrează în Figura 10.1). Odată ce a fost făcută legătura cu identitatea, nu e nici necesar, nici util să ne amintim de unde a venit informația, dat fiind că dezinformarea e agnostică în privința veridicității informației.

Compararea eficacității: informație, informație falsă și dezinformare — o perspectivă culturală asupra motivației bazate pe identitate

Convențiile limbajului îi fac pe oameni să presupună că emițătorul mesajului are în general scopul de a informa (principiul comunicării al lui Grice sau „logica comunicării“, Schwarz, 2014). Informarea presupune împărtășirea de conținut pe care persoana îl consideră adevărat, netendențios și potențial util pentru luarea unei decizii. Această logică se menține chiar dacă uneori emițătorii greșesc și, fără să știe, oferă informații false prin faptul că împărtășesc conținut neadevărat sau tendențios. Cu toate acestea, e posibil ca dezinformarea să aibă o putere mai mare pentru că se concentrează direct asupra modelării judecăților și manipulării comportamentelor. Astfel, informația falsă are o probabilitate mai mare de a influența acțiunea. Alte tehnici informaționale, cum ar fi construirea narativă, pot efectiv să schimbe opinii, ceea ce poate transforma schimbarea în acțiune dacă se face legătura cu identitățile sociale (Murphy, Frank, Moran și Patnoe-Woodley, 2011). Cu alte cuvinte, în loc să schimbe atitudini, tehnicile narrative pot lua calea persuasiunii centrate pe identitate (e.g. prin norme sociale, Paluck, 2009). Implicația este că tehnicile de persuasiune centrate pe identitate pot amplifica probabilitatea unei campanii de informare de a influența judecata și comportamentul, indiferent dacă informația este amintită sau nu. Provocarea cu care ne confruntăm în cazul corectării informațiilor false și dezinformării

constă în faptul că, odată ce întrebarea a fost cadrată drept „cum conceptualizăm noi asta“, este improbabil ca oamenii să treacă la o altă întrebare — „Care sunt dovezile care susțin această informație?“ Și mai rău este că, odată ce oamenii ajung să creadă că „noi“ gândim și ne purtăm într-un anumit mod, este mult mai improbabil să ia în considerare alte informații și să le considere orice altceva în afară de „fapte alternative“. Este improbabil să reușească încercările de corectare care nu sunt axate pe declanșarea unei procesări alternative bazate pe identitate. Sunt foarte necesare cercetări asupra cazurilor în care funcționează cadrele persuasive bazate pe identitate, asupra cazurilor în care nu funcționează și asupra modalităților de abordare a potențialului de abuz pe care îl au.

Notă

1. Această serie de argumente bazate pe informații, inclusiv statisticile referitoare la piața muncii, la prețurile bunurilor de consum și la profiturile din investițiile în UE, au fost în centrul principalei campanii de rămânere în UE, Britain Stronger in Europe, cu sloganul „Mai multe joburi, prețuri mai mici, drepturi pentru angajați“ (www.strongerin.co.uk/).

Referințe

Barberá, P., Jost, J., Nagler, J., Tucker, J. și Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26 (10), 1531–1542.

BBC. (2016). Interactive map with regional voter results June 23, 2016. EU

Referendum Results. Accesat la 9 septembrie 2019, www.bbc.com/news/politics/eu_referendum/results

Blinder, S. și Markaki, Y. (2018). Public attitudes toward EU mobility and Non-EU immigration: A distinction with little difference (No. WP-18–141). COMPAS, School of Anthropology, University of Oxford. Accesat la 9 septembrie 2019, www.compas.ox.ac.uk/2015/the-effects-of-immigration-on-nhs-waiting-times/

Boyd, R. și Richerson, P.J. (2005). The origin and evolution of cultures. New York, NY: Oxford University Press.

Brewer, M.B. (1979). In-group bias in the minimal intergroup situation: A cognitivemotivational analysis. *Psychological Bulletin*, 86 (2), 307–324.

Britain Stronger in Europe. (fără dată). Britain stronger in Europe. Accesat la 15 septembrie 2019, www.strongerin.co.uk/get_the_facts. Declarată drept campanie oficială de rămânere în UE de către Comisia Electorală la 13 aprilie 2016.

Bruner, J.S. și Postman, L. (1949). On the perception of incongruity: A paradigm. *Journal of Personality*, 18, 206–223.

Brush, S. și Weber, A. (2 iulie 2019). London's fight to remain a financial hub after Brexit. *Bloomberg*. Accesat la 9 septembrie 2019, www.bloomberg.com/news/articles/2019-07-03/london-s-fight-to-remain-a-financial-hub-after-brexit-quicktake

Cadwalladr, C. (18 martie 2018). „I made Steve Bannon's psychological warfare tool”: Meet the data war whistleblower. *The Guardian*. Accesat la 9 septembrie 2019, www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopherwyllie-faceook-nix-bannon-trump

Chiu, C.-Y., Gelfand, M., Harrington, J.R., Leung, A., Liu, Z., Morris, M., ... Zou, X. (2015). A conclusion, yet an opening to enriching the normative approach of culture. *Journal of Cross-Cultural Psychology*, 46 (10), 1361–1371.

Chiu, C.-Y., Gelfand, M., Yamagishi, T., Shteynberg, G. și Wan, C. (2010). Intersubjective culture: The role of intersubjective perceptions in cross-cultural research. *Perspectives on Psychological Science*, 5 (4), 482–493.

Cohen, D. (2001). Cultural variation: Considerations and implications. *Psychological Bulletin*, 127, 451–471.

Cooper, C. (26 iunie 2016). Brexit campaigners admit „there is no plan“ for what comes next as rivals plan Tory leadership bids. Accesat la 8 septembrie 2019, www.independent.co.uk/news/uk/politics/brexit-eu-referendum-campaigners-there-isno-plan-next-pm-tory-leadership-contest-a7104711.html

Douthat, R. (9 mai 2015). The suicide of Britain. *The New York Times*. Accesat la 9 septembrie

2019, www.nytimes.com/2015/05/10/opinion/sunday/ross-douthatthe-suicide-of-britain.html?searchResultPosition=4

Elmore, K. și Oyserman, D. (2012). If „we“ can succeed, „I“ can too: Identity-based motivation and gender in the classroom. *Contemporary Educational Psychology*, 37 (3), 176–185.

Facebook. (fără dată). Facebook advertising targeting options. Accesat la 15 aprilie 2019, www.facebook.com/business/ads/ad-targeting#

Facebook. (30 ianuarie 2019). Facebook reports fourth quarter and full year 2018 results. Accesat la 15 aprilie 2019, <https://investor.fb.com/investor-news/press-releasedetails/2019/Facebook-Reports-Fourth-Quarter-and-Full-Year-2018-Results/default.aspx>

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19 (4), 25–42.

Gilbert, D. (1991). How mental systems believe. *American Psychologist*, 46 (2), 107–119.

Gilbert, D., Krull, D. și Malone, P. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality*

and Social Psychology, 59 (4), 601–613.

Giuntella, O., Nicodemo, C., și Vargas-Silva, C. (2018). The effects of immigration on NHS waiting Times (No. Working Paper No. 124). COMPAS, School of Anthropology, University of Oxford. Accesat la www.compas.ox.ac.uk/2015/the-effects-ofimmigration-on-nhs-waiting-times/

Goodwin, M. (21 septembrie 2015). Why a „Brexit“ looms large. The New York Times. Accesat la 9 septembrie 2019, www.nytimes.com/2015/09/22/opinion/why-abrexit-looms-large.html?searchResultPosition=20

The Guardian. (2016). Interactive map with regional results 06/23/2016. EU Referendum: Full Results and Analysis. Accesat la 9 septembrie 2019, www.theguardian.com/politics/ng-interactive/2016/jun/23/eu-referendum-live-results-and-analysis?CMP=tw_twt_b-gdndata

Haberstroh, S., Oyserman, D., Schwarz, N., Kühnen, U. și Ji, L.J. (2002). Is the interdependent self more sensitive to question context than the independent self? *Journal of Experimental Social Psychology*, 38 (3), 323–329.

Haidle, M., Bolus, M., Collard, M., Conard, N., Garofoli, D., Lombard, M., ... Whiten, A. (2015). The nature of culture: An eight-grade model for the evolution and expansion of cultural capacities in hominins and other animals. *Journal of Anthropological Sciences*, 93, 43–70.

House of Commons, Digital, Culture, Media and Sport Committee, United Kingdom. (2018). Disinformation and „fake news“: Interim Report: Government Response to the Committee’s Fifth Report of Session 2017–19.

House of Commons, Digital, Culture, Media and Sport Committee, United Kingdom. (2019). Disinformation and „fake news“: Final Report (Eighth Report of Session 2017–19 ed.).

James, W. (1890). *The principles of psychology*. New York, NY: Dover.

Kahneman, D. și Frederick, S. (2002). Representativeness revisited: Attribute

substitution in intuitive judgment. În T. Gilovich, D. Griffin, & D. Kahneman (ed.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). New York, NY: Cambridge University Press.

Kang, C. și Frenkel, S. (4 aprilie 2018). Facebook says Cambridge Analytica harvested data of up to 87 million users. *The New York Times*. Accesat la 9 septembrie, 2019, www.nytimes.com/2018/04/04/technology/mark-zuckerberg-testify-congress.html

Kosinski, M., Stillwell, D. și Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110 (15), 5802–5805.

Kramer, A., Guillory, J. și Hancock, J. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111 (24), 8788–8790.

Kurzban, R. și Neuberg, S. (2005). Managing ingroup and outgroup relationships. În D. Buss (ed.), *The handbook of evolutionary psychology* (pp. 653–675). Hoboken, NJ: John Wiley & Sons.

Lewis, N., Jr. și Oyserman, D. (2015). When does the future begin? Time metrics matter, connecting present and future selves. *Psychological Science*, 26 (6), 816–825.

Massie, A. (9 mai 2015). Reflections on the revolution in Scotland. *The Spectator*. Accesat la 9 septembrie 2019, www.spectator.co.uk/2015/05/this-election-hasbeen-all-about-scotland/

McCrae, R. (2018, July 31). Geographical breakdown of the current account. *The Pink Book*. Accesat la 9 septembrie 2019, www.ons.gov.uk/economy/nationalaccounts/balanceofpayments/datasets/9gec

Mourey, J., Lam, B. și Oyserman, D. (2015). Consequences of cultural fluency. *Social Cognition*, 33 (4), 308–344.

Mourey, J., Oyserman, D. și Yoon, C. (2013). One without the other: Seeing relationships in everyday objects. *Psychological Science*, 24 (9), 1615–1622.

Murphy, S., Frank, L., Moran, M. și Patnoe-Woodley, P. (2011). Involved, transported, or emotional? Exploring the determinants of change in knowledge, attitudes, and behavior in entertainment-education. *Journal of Communication*, 61 (3), 407–431.

Nisbett, R. și Norenzayan, A. (2002). Culture and cognition. În D. Medin (coord.), *Stevens' handbook of experimental psychology* (pp. 561–597). New York, NY: Wiley.

Nurra, C. și Oyserman, D. (2018). From future self to current action: An identity-based motivation perspective. *Self and Identity*, 17 (3), 343–364.

Oyserman, D. (2007). Social identity and self-regulation. În A. Kruglanski și E.T. Higgins (coord.), *Social psychology: Handbook of basic principles* (pp. 432–453). New York, NY: Guilford Press.

Oyserman, D. (2009). Identity-based motivation: Implications for action-readiness, procedural-readiness, and consumer behavior. *Journal of Consumer Psychology*, 19, 250–260.

Oyserman, D. (2011). Culture as situated cognition: Cultural mindsets, cultural fluency, and meaning making. *European Review of Social Psychology*, 22 (1), 164–214.

Oyserman, D. (2015a). Culture as situated cognition. În R. Scott și S. Kosslyn (coord.), *Emerging trends in the behavioral and social sciences* (pp. 1–11). Hoboken, NJ: John Wiley and Sons.

Oyserman, D. (2015b). *Pathways to success through identity-based motivation*. New York, NY: Oxford University Press.

Oyserman, D. (2017). Culture three ways: Culture and subcultures within countries. *Annual Review of Psychology*, 68, 435–463.

Oyserman, D. (2019a). Cultural fluency, mindlessness, and gullibility. În R. Baumeister și J. Forgas (ed.), *The social psychology of gullibility*. New York, NY: Routledge.

Oyserman, D. (2019b). The essentialized self: Implications for motivation and selfregulation. *Journal of Consumer Psychology*, 29 (2), 336–343.

Oyserman, D., Elmore, K., și Smith, G. (2012). Self, self-concept, and identity. În M. Leary și J. Tangney (coord.), *Handbook of self and identity* (pp. 69–104). New York, NY: Guilford Press.

Oyserman, D., Lewis, N., Jr., Yan, V., Fisher, O., O'Donnell, S.C. și Horowitz, E. (2017). An identity-based motivation framework for self-regulation. *Psychological Inquiry*, 28, 139–147.

Oyserman, D., Sorensen, N., Reber, R. și Chen, S. (2009). Connecting and separating mind-sets: Culture as situated cognition. *Journal of Personality and Social Psychology*, 97 (2), 217–235.

Oyserman, D. și Yan, V.X. (2019). Making meaning: A culture-as-situated cognition approach to the consequences of cultural fluency and disfluency. În S. Kitayama și D. Cohen (ed.), *Handbook of cultural psychology* (pp. 536–565). New York, NY: Guilford Press.

Petty, R. și Cacioppo, J.T. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Dubuque, IA: Wm. C Brown.

Petty, R. și Cacioppo, J.T. (1984). Source factors and the elaboration likelihood model of persuasion. *Advances in Consumer Research*, 11, 668–672.

Paluck, E.L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology*, 96 (3), 574–587.

Platow, M., Foddy, M., Yamagishi, T., Lim, L. și Chow, A. (2012). Two experimental tests of trust in in-group strangers: The moderating role of common knowledge of group membership. *European Journal of Social Psychology*, 42 (1), 30–35.

Redazione. (9 mai 2018). Exclusive interview with Christopher Wylie, the Cambridge Analytica whistleblower. *Vogue Italia*. Accesat la 9 septembrie

2019, www.vogue.it/en/news/daily-news/2018/05/09/interview-with-christopher-wylie-cambridgeanalytica/

Reyna, V. (2004). How people make decisions that involve risk: A dual-processes approach. *Current Directions in Psychological Science*, 13 (2), 60–66.

Schwarz, N. (2014). *Cognition and communication: Judgmental biases, research methods, and the logic of conversation*. New York, NY: Psychology Press.

Schwarz, N., Strack, F., Hilton, D. și Naderer, G. (1991). Base rates, representativeness, and the logic of conversation: The contextual relevance of „irrelevant“ information. *Social Cognition*, 9 (1), 67–84.

Schwarz, N. și Sudman, S. (coord.). (2012). *Context effects in social and psychological research*. Berlin and Heidelberg, Germania: Springer Science & Business Media.

Shochat, E., Shniberg, M., Hirsch, G. și Tagiman, Y. (2009). U.S. Patent No. US8666198B2. Washington, DC: U.S. Patent & Trademark Office. Relationship mapping employing multi-dimensional context including facial recognition.

Shteynberg, G. (2015). Shared attention. *Perspectives on Psychological Science*, 10 (5), 579–590.

Shweder, R. și LeVine, R. (ed.). (1984). *Culture theory: Essays on mind, self and emotion*. New York, NY: Cambridge University Press.

Sudman, S., Bradburn, N.M. și Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. San Francisco, CA: Jossey-Bass.

Triandis, H. (2007). Culture and psychology. În S. Kitayama și D. Cohen (coord.), *Handbook of cultural psychology* (pp. 59–76). New York, NY: Guilford Press.

UK Research and Innovation (9 august 2019) Government pledges to protect science and research post Brexit [comunicat de presă]. Accesat la www.gov.uk/government/news/government-pledges-to-protect-science-and-research-post-brexit?utm_source=3aa9d710-efe1-4eea-b45f-043b81d05c9d

Weedon, J., Nuland, W. și Stamos, A. (2017). Information operations and Facebook. Accesat la 9 septembrie 2019, <https://fbnewsroomus.files.wordpress.com/2017/04/facebookand-information-operations-v1.pdf>.

Capitolul 11. Convingeri conspiraționiste

Cunoaștere, apărări ale eului și integrare socială în procesarea fake news

Dolores Albarracín

Pesemne nu e greu să dai de fake news, fiindcă nu mi-a luat decât șase minute să găsesc acest text:

Au trecut 15 luni de când unul dintre principalii oameni de știință de la CDC [Centrul pentru Prevenția și Controlul Bolilor], dr. William Thompson, a recunoscut că un studiu realizat în 2004 de CDC a fost falsificat pentru a arăta că nu există o legătură între vaccinul ROR și autism. În august 2014, dr. Thompson a spus: „Regret că eu și coautorii mei am omis informații semnificative statistic în articolul nostru din 2004, publicat în Journal of Pediatrics.“

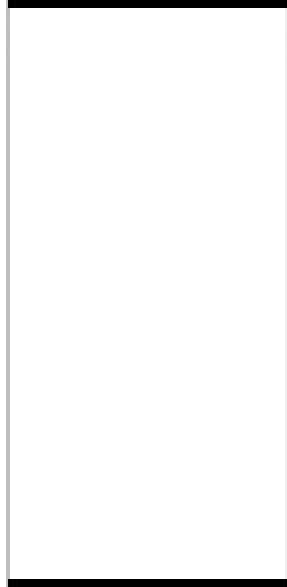
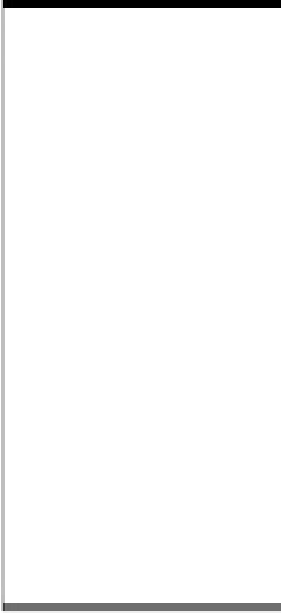
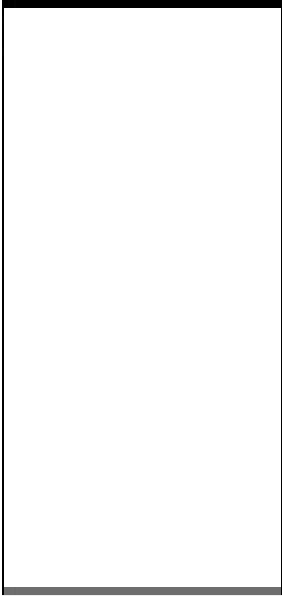
Oameni buni, un cercetător de la CDC a recunoscut sub jurământ că CDC a falsificat date pentru a nega legătura dintre vaccinare și autism. Mai mult, dr. Thompson a afirmat că cercetători seniori de la CDC au încercat să distrugă toate documentele care aveau legătură cu această mușamalizare. Dr. Thompson a salvat documentele. Au trecut aproape 15 luni și nu s-a întâmplat nimic.

(<https://healthimpactnews.com/2015/cdc-cover-up-of-autism-and-vaccine-link-continues/>)

Chiar ar trebui să credem că există dovezi că vaccinul ROR (anti-rujeolă, oreion, rubeolă) provoacă autism și că s-a încercat ca aceste dovezi să fie ascunse? Convingeri conspiraționiste precum cele promovate în „fake news” creează o legătură dublă (double bind) care ne spune și (1) să credem informația pe baza dovezilor, și (2) să ne mulțumim cu verificări imperfecte, pentru că dovezile au fost ascunse. Prin această dublă legătură, teoriile conspirației propun o pseudorealitate care este plauzibilă sau aproape științifică și, în același timp, își păstrează aura de gândire magică și anticipare a faptului că s-ar putea ca adevărul să nu iasă niciodată la iveală. Modelul pe care îl prezint în acest capitol descrie modul în care astfel de argumente conspiraționiste și alte tipuri de informații false par plauzibile și nefalsificabile pe baza necesității de cunoaștere, apărării eului și integrării sociale. Motivația bazată pe cunoaștere reprezintă nevoia de a avea reprezentări acurate și complete ale lumii. Motivația determinată de apărările eului reprezintă mecanismele defensive prin care sinele este protejat de emoții sau sentimente negative, ca frica. Motivația integrării sociale reprezintă necesitatea de a ne înțelege și de a fi valorizați de ceilalți.

Mai întâi clasific convingerile în funcție de verificabilitate și plasez convingerile conspiraționiste și teoriile conspirației pe acest spectru al verificabilității. Expun apoi modul în care teoriile conspirației și alte lucruri pe care le susțin fake news se leagă de (1) necesitatea de cunoaștere, (2) necesitatea de apărare a eului și (3) necesitatea de integrare socială. Dacă este implicată doar una dintre aceste motivații, probabil că astfel de convingeri nu ar fi atât de seducătoare. Însă teoriile conspirației le permit receptorilor de știri „să aibă și cioara din mână, și vrabia de pe gard”, prin faptul că îi fac să se simtă „informați” și în același timp să se complacă în minciuni convenabile care le menajează orgoliul sau care îi mențin integrați social cu prețul lipsei acurateții. Cadrul teoretic al acestui capitol este ilustrat în Figura 11.1.

Figura 11.1. Spectrul verificabilității și motivațiile umane



Motivație bazată pe cunoaștere

Verificabil Neverificabil

Motivație bazată pe apărări ale eului

Motivație bazată pe integrarea socială

Verificabilitatea convingerilor

Convingerile pot fi (1) verificabile sau (2) neverificabile. Zona verificabilității spectrului include propozițiile bazate pe observație directă prin intermediul simțurilor și pe metoda științifică. Grupurile de oameni de știință, de exemplu, sunt constituite pe baza unei metode explicite de verificare cu care toți sunt de acord. Jurnalismul de investigație are principii similare în privința domeniului politic, educațional sau financiar. În sfera convingerilor verificabile, convingerile eronate sunt convingeri despre care se pretinde că sunt corecte, dar care nu au corespondent în realitate (Chan, Jones și Albarracín, 2017). Numeroase convingeri eronate sunt în centrul grupurilor sociale (e.g. „antivacciniști“) care le susțin. O teorie este o serie organizată de ipoteze pe baza cărora se fac predicții într-un mod sistematic despre lucruri din realitate. Teoriile științifice folosesc metoda observației sistematice și obțin reproduceri, compară condiții și dezbat deschis concluziile.

Zona neverificabilității include (1) teoriile conspirației și (2) teorii religioase care implică forme nesistematice și indirecte de culegere de dovezi. O teorie a conspirației de tipul celor care apar din plin în fake news este o structură de convingeri conform cărora forțe puternice și ascunse sunt responsabile de evenimente ale căror cauze le-au ascuns (Douglas, Sutton, Jolley și Wood, 2015; Enster, 1999; Groh, 1987; Pratt, 2003; van Prooijen și Krouwel, 2015). Teoriile conspirației au modalități de a inocula membrii împotriva atacurilor potențiale, în speță prin inculcarea neîncrederii în informațiile care provin din

surse externe. Aceste structuri de convingeri sunt solide și rezistente la verificare.

Convingerile religioase sunt susținute de o entitate supranaturală și pot avea fie o structură solidă, fie o structură slabă. Sistemele slabe sunt compuse din convingeri interrelaționate, dar în general sunt mai simple și/sau mai puțin organizate și nu au mecanisme de a bloca informațiile care le contrazic. De exemplu, universalismul desemnează mai multe convingeri de tip „new age”, credințe păgâne și religii tribale conform cărora zeitatea este întregul univers sau Pământul. Universalismul include și sisteme în care se crede că Dumnezeu este în toate lucrurile sau în toți oamenii. Sistemele universaliste sunt în general neconturate și nu au un cod scris de comportament și nici nu dispun că toate convingerile divergente sunt greșite.

Sistemele religioase puternice, cum ar fi sectele, sunt ramificații ale unui sistem religios mai amplu și pun accentul pe legitimitatea unică a convingerilor și practicilor grupului (Meagher și O'Brien, 1979). Sectele mențin granițe ferme cu restul societății și cu grupul mai mare din care s-au separat. Grupul Boko Haram din Nigeria este un exemplu de sectă, întrucât se definește ca grup disident și respinge educația tradițională în favoarea secretului și a cunoașterii privilegiate. În al doilea rând, fundamentalismul presupune aderența strictă la litera textelor religioase. De exemplu, Asociația Evanghelică a lui Billy Graham susține o interpretare literală a Bibliei (e.g. Geneza) prin intermediul unui imperiu mass-media. Extremismul islamic este, desigur, bine cunoscut. Alte exemple sunt kharigiții din secolul al VII-lea și fundamentalismul wahhabi care se presupune că este susținut de Arabia Saudită, Pakistan și Qatar (Braswell, 2000; Delong-Bas, 2008; Huff, 1995; Nasir, 2009). Toate aceste sisteme de convingeri sunt solide sub aspectul neverificabilității, pentru că implică adevăruri revelate și în același timp descurajează adepții de la a căuta schimbarea.

Spectrul verificabilității și motivațiile umane

Spectrul verificabilității apare în Figura 11.1 și oferă un cadru interesant pentru înțelegerea modului în care convingerile pot servi diferite tipuri de motivații. În capătul din stânga, convingerile cele mai verificabile servesc în primul rând o funcție de cunoaștere, dat fiind că pot fi aliniate realității. În capătul din dreapta, cele mai puțin verificabile convingeri pot servi în primul rând funcției sociale de integrare, dat fiind că avem nevoie să auzim de la alții povești despre lumi secrete sau spirituale. În centru, și convingerile verificabile, și cele neverificabile servesc funcții de apărare a eului, cele verificabile fiind în slujba funcțiilor de acțiune și control (i.e. nevoia de a acționa eficient și de a controla rezultatul), cele neverificabile, prin funcțiile exclusiv defensive ale eului. Convingerile verificabile susțin acțiunea și controlul în interacțiunea cu lucruri din realitate, iar cele neverificabile ne susțin iluziile convenabile.

Date fiind considerațiile pe care le-am făcut în legătură cu Figura 11.1, teoriile conspirației pot servi oricăreia dintre aceste funcții și se află, după cum am sugerat, în partea centrală a spectrului, între verificabilitate și neverificabilitate. Sunt suficient de aproape de verificabilitate pentru a fi plauzibile și sunt totodată îndeajuns de nefalsificabile pentru a fi neverificabile. Într-o analiză a tweet-urilor referitoare la vaccinuri, noi (Palmer și Albaraccín, 2018) am clasificat tweet-urile referitoare la ROR și HPV în termeni de anti și pro-vaccinare. Rezultatele au indicat prevalența conținutului științific, cu un nivel puțin mai scăzut de științificitate în postările anti-vaccinare în comparație cu cele pro-vaccinare. Aceste date sunt prezentate în Tabelul 11.1 și indică faptul că unele teorii ale conspirației sunt susținute într-un mod care le face să pară plauzibile din punct de vedere științific, așadar mai aproape de capătul verificabilității al spectrului din Figura 11.1. Însă, în același timp, prevalența conținuturilor care denunță o mușamalizare demonstrează neverificabilitatea conținuturilor anti-vaccinare, întrucât credința că a avut loc o mușamalizare anulează capacitatea audienței de a judeca informația. Subterfugiul constă în plasarea mesajelor conspiraționiste într-o zonă gri, astfel încât să fie și verificabile, și neverificabile în același timp.

Tabelul 11.1. Aspecte ale tweet-urilor referitoare la vaccinuri

Aspect	vaccin ROR N = 857 vaccin HPV N = 737		
	Anti	Pro	Al
Conținut științific	16%	18%	14
Denunțare a unei mușamalizări	18%	0%	8%

Am cules și alte date care arată că teoriile conspirației sunt diferite de adevăratele conspirații și în privința plauzibilității, și a nefalsificabilității. Într-o serie de studii, le-am cerut participanților să evalueze teorii, unele adevărate, altele false, evaluând convingerile pe care le implică, precum și plauzibilitatea și nefalsificabilitatea percepută a fiecăreia. Printre aceste teorii au fost:

(1) Barack Obama nu s-a născut în SUA; și-a falsificat certificatul de naștere pentru a deveni președinte.

(Politico, 2011)

(2) Imigranții ilegali care au votat ilegal au împiedicat Partidul Republican să câștige votul popular în 2016.

(Business Insider, 2018)

(3) Guvernul SUA a creat epidemia de HIV prin injectarea experimentală a virusului în oameni cu ascendență africană.

(Heller, 2015)

(4) Vaccinul ROR cauzează autism, dar acest fapt a fost mușamalizat de guvernul SUA.

(Eggertson, 2010)

(5) Hillary Clinton a condus o pizzerie care funcționa ca acoperire pentru

prostituție infantilă în Washington, D.C.

(Rolling Stone, 2017)

(6) Hibrizi de oameni și extraterestri reptilienii ocupă în prezent poziții de putere.

(goo.gl.T79SzH)

Fiecareia dintre aceste teorii îi corespundea una adevărată. De exemplu, teoriei pizzeriei îi corespundea ca eveniment adevărat faptul că Hillary Clinton își folosea propriul server pentru comunicări secrete. În general, deși a existat o mare variabilitate a convingerilor în teoriile prezentate, participanții au avut mai mare încredere în teoriile adevărate decât în cele false. Mai mult, am putut să studiem corelațiile dintre convingeri și plauzibilitate și nefalsificabilitate. Plauzibilitatea a fost definită ca înțelegere a motivului pentru care fiecare tip de situație ar fi putut avea loc și amintire a altor exemple din trecut de situații asemănătoare. Nefalsificabilitatea a fost definită drept convingere că dovezile contradictorii sunt rezultatul unei mușamalizări sau că sunt false. Analizele au arătat în general asociații sănătoase ale convingerilor cu fiecare dimensiune, deși nefalsificabilitatea a corelat pozitiv în mai mare măsură cu credința în teorii conspiraționiste decât cu alte credințe. Conform predicției pe baza faptului că o credință conspiraționistă este parțial verificabilă și parțial neverificabilă, ambele dimensiuni au fost importante.

În secțiunile următoare, explic modul în care teoriile conspirației sunt susținute de motivații determinate de cunoaștere, de apărările eului și de motive de integrare socială. Să ne gândim mai întâi la importanța motivației de cunoaștere. Construim teorii pentru că sunt explicații utile ale observațiilor pe care le facem asupra realității. Stelele au exercitat o atracție masivă asupra nespecialiștilor, filosofilor și oamenilor de știință, toți încercând să explice observații deconcertante: obiecte astronomice care dispăreau, care deveneau mai mari, care își schimbau traiectoria etc. „Mișcarea retrogradă“, de

exemplu, este observația că uneori planetele par să se miște invers pe firmament. Ptolemeu a trebuit să creeze o explicație elaborată a mișcării retrograde, pentru că modelul său susținea că plantele și stelele se mișcă înainte în jurul Pământului. Explicația lui implica un aranjament complicat de cercuri, în care o planetă se rotea în jurul Pământului pe o orbită circulară. Însă, la un moment dat, planeta intra într-o mică buclă înainte de a-și relua traiectoria pe orbita principală. Intrând în această buclă, planeta, care se mișca întotdeauna înainte, „părea“ că se mișcă înapoi dacă era observată de pe Pământ.

Modele explicative complicate, precum cel al lui Ptolemeu, sunt adesea abandonate în favoarea unei explicații mai cumpătate. În cazul mișcării planetelor este explicația oferită de Copernic pentru mișcarea retrogradă. El a revoluționat cunoașterea științifică și religioasă a universului prin propunerea ipotezei că Pământul orbitează în jurul soarelui. Viteza de rotație și orbitele sunt însă diferite, ceea ce creează iluzia de mișcare retrogradă când Pământul se mișcă cu o viteză mai mare decât celelalte corpuri cerești. În acest caz, o explicație mai simplă a înlocuit-o pe cea anterioară.

Desigur, teoria lui Ptolemeu este o explicație întru totul utilă a mișcării corpurilor cerești. Pentru explicarea motivului înlocuirii unei astfel de teorii utile și larg acceptate social este necesar conceptul de motivație epistemică (Kruglanski, 1989, 1990; Kruglanski, Dechesne, Orehek și Pierro, 2009; Kruglanski, Orehek, Dechesne și Pierro, 2010), numit de alți autori motivație a acurateții (Hart et al., 2009). Acest concept se referă la faptul că oamenii caută cunoașterea de dragul cunoașterii, chiar dacă nu are aplicații practice. Motivația cunoașterii este și baza convingerilor întemeiate, și a celor eronate. Motivația cunoașterii pare o sintagmă mai potrivită decât motivație a acurateții, dat fiind că informația și cunoașterea se pot plasa oriunde pe spectru, de la lipsa de acuratețe la acurate. William Davies (2016), autor premiat de New York Times, a intitulat o descriere a campaniei prezidențiale din SUA din 2016 „Epoca politicii post-adevăr“. Potrivit lui Davies, tendința de a reprezenta societatea în termeni de fapte a început odată cu introducerea contabilității la finalul Evului Mediu. Însă, până atunci, faptele nu erau la fel de importante, cunoașterea și adevărul fiind dezvăluite într-o manieră nonliterală, înlocuind sfera mai degrabă trivială a faptelor.

Pe lângă cunoaștere, suntem impulsionați de motivația de apărare a Eului, adică protejarea sinelui de sentimente neplăcute, cum ar fi dezamăgirea, rușinea sau tristețea care pot fi provocate de comportarea în moduri incompetente sau reproșabile din punct de vedere social. O vastă literatură centrată asupra autoperfecționării și raționamentului motivat***** sugerează că oamenii își modelează convingerile în funcție de necesitatea psihologică de creștere a stimei de sine (Baumeister, 1997; Eagly și Chaiken, 1993; Kunda, 1990; Tesser, 2001). De exemplu, se apără respingând sau modificând informațiile care diminuează valoarea percepută a sinelui. Dar eul se ocupă și de acțiuni și de adaptarea la realitate. Din această perspectivă, motivația de a acționa echilibrează și motivația de a cunoaște, ca în cazul în care suntem obligați să luăm decizii pe baza unor informații în mod necesar incomplete.

Necesitățile de cunoaștere și de apărare a eului activează obiective care pun în mișcare mai multe procese psihologice esențiale, inclusiv centrarea asupra anumitor concepte, utilizarea unor instrumente rudimentare de raționare, ca euristicele și atitudinile, și aplicarea unei analize mai laborioase a informației. Iar trecerea prin aceste etape depinde de nivelul de încredere pe care oamenii îl caută sau de ceea ce Shelly Chaiken (vezi Chaiken și Mahswaran, 1994; Chen și Chaiken, 1999; Chen, Duckworth și Chaiken, 1999) a numit „praguri de încredere“. Un nivel relativ scăzut al nivelului dorit de încredere ar duce la activarea conceptului și la procese rudimentare de raționare. Dorința de a avea un nivel mai mare de încredere îi va face pe oameni să aplice un tip de raționare analitică mai profundă până când nivelul concret de încredere ajunge la pragul dorit.

Pragurile de încredere sunt importante și pentru motivația de cunoaștere, și pentru motivația de apărare a eului. Oamenii care încearcă să ajungă la o înțelegere profundă a unei probleme vor aplica mai întâi activarea conceptului și vor încerca să apeleze la un raționament de tip euristic. Dacă euristica le satisface nivelul de încredere dorit, este probabil să ajungă la o judecată bazată pe aceasta. Invers, dacă nu se ajunge la nivelul dorit de încredere, este probabil ca oamenii să continue printr-o analiză a informației care implică efort și care se încheie când este atins pragul dorit de încredere. Chiar dacă aceste mecanisme sunt determinate și de motivația cunoașterii, și de cea a apărării eului, procesarea detaliată are efecte diferite, în funcție de

tipul motivației. Dacă este căutată cunoașterea, procesele superioare duc la o înțelegere mai detaliată, nedistorsionată. Cu toate acestea, dacă este căutată apărarea eului, procesarea superioară pur și simplu raționalizează o concluzie dorită.

Oamenii se străduiesc, de asemenea, să se integreze social, iar această necesitate influențează la rândul său convingerile pe care le au. De exemplu, convingerea că strămoșii noștri ne apără din viața de apoi ne poate susține eul, dar ne poate oferi și o formă spirituală de integrare socială. Mai mult, numeroase teorii ale conspirației și majoritatea convingerilor științifice creează automat legături sociale cu o comunitate de semeni care împărtășesc aceste credințe.

Teorii ale conspirației și nevoia de cunoaștere

Înțelegerea realității presupune construirea cunoașterii și realizarea de judecăți inerent subiective. Cunoașterea este ceea ce credem că există independent de noi înșine: scriem cu certitudinea că tastatura din fața noastră există și suntem la fel de siguri că avem nevoie de parola corectă pentru a avea acces la computer. Pe lângă această cunoaștere, putem face judecăți, conștienți fiind că sunt subiective. Exemple sunt credința în Dumnezeu sau convingerea că renunțarea la fumat are un efect pozitiv pentru starea de sănătate. Aceste convingeri poziționează fiecare persoană în raport cu altele, dar nu descriu realitatea sau probabilitatea absolută ca o propoziție să fie adevărată.

Cunoașterea, credințele și atitudinile joacă mai multe roluri esențiale, inclusiv de reprezentare a realității, de înțelegere a aspectelor obiective ale lumii și de comparație a acestor aspecte cu standarde subiective. De exemplu, un tâmplar poate măsura blatul mesei pe care urmează să o creeze ca având un diametru de un metru. Însă, dacă acest diametru înseamnă o masă mare sau mică nu se

poate evalua prin măsurătoare, ci implică prezența unui observator care are un standard mental pe baza căruia face o judecată, probabil diferită de a altora. Firește, reprezentările corecte ale realității ne permit să navigăm prin lumea exterioară. În acest context, avantajul de a avea convingeri este mai puțin clar și ne aduce într-un punct important: creăm convingeri când trebuie să facem predicții în privința a ceva necunoscut pentru a îndeplini un obiectiv („Cred că Mary se va duce la conferință, așa că o să încerc să-mi fac timp să vorbesc cu ea“). Căutăm o aproximare care ne va permite să acționăm în încercarea de a îndeplini obiective, recunoscând că nu avem cunoștințe precise pentru fiecare caz în parte. Atitudinile ne permit să înțelegem că o pisică nu reprezintă aceeași amenințare ca un leu, că anumite obiecte sunt comestibile și altele, nu și că unii oameni ne vor recompensa, iar alții ne vor pedepsi.

Dar teoriile conspirației facilitează și înțelegerea pe care o avem asupra lumii? Da, un volum considerabil de cercetări sugerează că așa stau lucrurile. Există trei predictorii ai teoriilor conspirației care indică această concluzie, și anume (1) tendința de a percepe tipare, (2) incertitudinea și (3) necesitatea de reducere a incertitudinii.

Tendința de a percepe tipare

Un aspect important al teoriilor conspirației este că fac legături între puncte de pe o hartă care nu există. Astfel, cercetările au acordat atenție unor factori precum tendința de a percepe tipare. De exemplu, cercetarea realizată de Moulding et al. (2016) a inclus diferite evaluări ale teoriilor conspirației, ca Intolerance for Uncertainty Scale (Scala de intoleranță la incertitudine) (Buhr și Dugas, 2002) și Need for Cognitive Closure (Necesitatea de încheiere cognitivă) (Webster și Kuglanski, 1994). Acest studiu a evaluat și tendința de a percepe un tipar în relații sociale și în distribuirea justă a durerii. Specific, autorii au inclus World Assumptions Scale (Scala presupunerilor despre lume) (WAS; Janoff-Bulman, 1989), care, printre alte dimensiuni, evaluează

percepția întâmplărilor aleatorii („oamenilor li se întâmplă lucruri rele în mod aleatoriu”). Evaluările credinței în teorii ale conspirației au inclus Belief in Conspiracy Theories Inventory (Inventarul credinței în teorii ale conspirației) (Swami, Chamorro-Premuzic și Furnham, 2010). Această scală folosește cote de la 1 (complet fals) la 10 (complet adevărat) pentru itemi ca „Aselenizările Apollo nu au avut loc niciodată și au fost contrafăcute într-un studio de film de la Hollywood”. Studiul a inclus și alte evaluări ale convingerilor conspiraționiste, toate având o corelație pozitivă moderată cu nevoia de încheiere cognitivă și cu tendința de a percepe tipare. Cercetări mai recente au ajuns la concluzii similare (Van der Wal, Sutton, Lange și Braga, 2018; van Prooijen, Douglas și De Inocencio, 2018).

Incertitudinea

Teoriile conspirației servesc o funcție de cunoaștere, dat fiind că le permit oamenilor să perceapă lumea ca fiind organizată și predictibilă (Bale, 2007; Clarke, 2002; Hofstadter, 1964; Miller, 2002; van Prooijen și Jostmann, 2013). Deși orice formă de explicație servește funcția cunoașterii, teoriile conspirației excelează datorită structurii lor de sistem monolitic. Acest tip de structură, pe care eu o numesc „sistem puternic”, face nefalsificabile o serie de explicații prin faptul că pune la îndoială orice dovadă împotriva teoriei. De exemplu, teoria că reptile extraterestre ocupă poziții de putere pe Pământ conține mecanisme de invalidare a dovezilor — extraterestrii au capacitatea de a se metamorfoza și de a ascunde orice dovadă a propriei existențe. Acest mecanism aduce suspiciuni în privința oricărei informații oferite de persoane care doresc să o contrazică și prioritizează astfel validitatea perspectivei conspiraționiste.

Din punct de vedere psihologic, incertitudinea este definită ca sentiment de îndoială, care în general se presupune că e cel puțin ușor neplăcut (Wichman et al., 2010). Incertitudinea este provocată de o discrepanță între nivelul concret de înțelegere a unui eveniment sau fenomen și nivelul dorit de

înțelegere (Park, 2010). Cercetări realizate de van Prooijen și Jostmann (2013) au arătat modul în care incertitudinea susține credința în teorii ale conspirației fără a afecta alte convingeri. În cadrul primului experiment expus în această lucrare, participanților le-a fost indus sentimentul incertitudinii prin sarcina de a scrie despre o experiență de incertitudine sau li s-a cerut să scrie despre experiența de a se uita la televizor. Apoi au citit unul sau două fragmente, primul, despre politicile neetice de resurse umane ale companiilor petroliere (etichetate „imorale“ în articol) și celălalt, despre politici prezentate ca fiind etice (etichetate „morale“ în articol). Chiar dacă aceste fragmente nu discutau specific o conspirație, este întemeiată așteptarea ca un actor descris ca imoral să reprezinte terenul ideal pe care o teorie a conspirației poate completa spațiile goale. Astfel, după ce au citit informațiile despre companiile petroliere, participanților li s-a cerut să răspundă la trei întrebări: „Credeți că trusturile petroliere aveau un interes direct în războiul din Irak?“, „Credeți că trusturile petroliere au contribuit la declanșarea războiului din Irak?“ și „În ce măsură credeți că oameni asociați cu companii petroliere au dat ordinul de începere a războiului din Irak?“ Răspunsurile la acești itemi au avut un nivel ridicat de validitate internă și au fost considerate o măsură a credinței în teoriile conspirației.

Ipoteza care a ghidat primul experiment realizat de van Prooijen și Jostmann (2013) a fost că stările de incertitudine vor duce la credința spontană în conspirații în cazul actorilor neetici (vs. etici), chiar dacă nu sunt prezentate informații conspiraționiste. Cu toate astea, în absența incertitudinii, cercetătorii se așteptau să nu existe nicio diferență între condiția neetică și cea etică. În acest experiment, participanții au avut o probabilitate mai mare de a deduce o conspirație dacă un actor era perceput ca fiind în general neetic (vs. etic), dar acest efect a fost prezent doar când incertitudinea a fost făcută accesibilă.

Rezultatele Experimentului 1 au fost reproduse într-un al doilea experiment, cu aceeași manipulare a incertitudinii, dar folosindu-se o descriere a autorităților dintr-un oraș african ca fiind corupte sau ca respectând legea. După citirea acestei descrieri, participanții au aflat că un candidat la alegerile prezidențiale se opunea acestor autorități și a murit într-un accident de mașină. Și în acest caz, dacă oamenii care resimt incertitudinea construiesc o teorie conspiraționistă ca modalitate de a reduce incertitudinea, actorii neetici

ar trebui să ducă la o mai mare acceptare a întreprinderilor conspiraționiste decât actorii etici. După cum arată rezultatele, chiar așa s-a întâmplat.

Mai multe dovezi indirecte ale efectului incertitudinii provin din cercetări asupra efectelor ambivalenței asupra susținerii convingerilor conspiraționiste. Van Harreveld, Rutjens, Schneider, Nohlen și Keskinis (2014) au manipulat ambivalența cerându-le participanților să descrie o chestiune sau o opinie în legătură cu care erau ambivalenți sau asupra căreia aveau o perspectivă neechivocă. Participanții au evaluat apoi gradul în care rezultatele financiare descrise într-o vinietă aveau legătură cu acțiunile altcuiva. Rezultatele au indicat că manipularea ambivalenței a provocat ambivalență și afecte negative în legătură cu ambivalența. Mai mult, sentimentele neplăcute asociate îndoielii au amplificat la rândul lor credințele conspiraționiste.

Necesitatea de a reduce incertitudinea

De una singură, incertitudinea ar avea un efect redus dacă oamenii s-ar mulțumi fără certitudini, și există cu siguranță diferențe în ce privește necesitatea indivizilor de a reduce incertitudinea. Așadar, deloc surprinzător, mai multe studii au examinat relația dintre aceste diferențe individuale și convingerile eronate. Necesitatea de a reduce incertitudinea este adesea măsurată cu Intolerance of Uncertainty Scale (Carleton, Norton și Asmundson, 2007; Freeston, Rhéaume, Letarte, Dugas și Ladouceur, 1994), prin afirmații ca „Nu suport să fiu luat prin surprindere” cotate pe o scală Likert de cinci puncte (de la 1 = nu mă caracterizează deloc la 5 = mă caracterizează întru totul). O altă măsurătoare folosită frecvent este Need for Cognitive Closure Scale (Roets și Van Hiel, 2011; Webster și Kruglanski, 1994), cu afirmații ca „Nu-mi plac situațiile imprevizibile”, evaluate pe o scală de tip Likert de șase puncte (de la 1 = total dezacord la 6 = total de acord). Există o scală foarte similară acesteia, Uncertainty Orientation Scale (Scala orientării în situații incerte) (Sorrentino și Short, 1986), care se pare că a fost rezultatul unei strânse colaborări între echipe care în cele din urmă s-au

separat. S-ar putea prezice că oricare dintre aceste evaluări indică o asociație pozitivă cu teoriile conspirației.

Însă, contrar așteptărilor, datele științifice au adus dovezi slabe în favoarea unei asociații directe între teoriile conspirației și necesitatea de a reduce incertitudinea. Să ne amintim că studiul realizat de van Prooijen și Jostmann (2013) a inclus cinci măsurători ale susținerii teoriilor conspirației. Cu o astfel de meticulozitate, studiul oferă date excelente pentru determinarea gradului în care convingerile conspiraționiste corelează cu necesitatea de finalizare cognitivă și cu intoleranța incertitudinii. Cu toate acestea, datele au fost dezamăgitoare, fiind semnificativă doar asociația dintre rezultatele la Conspiracy Mentality Questionnaire (Chestionarul de mentalitate conspiraționistă) și Intolerance for Uncertainty.

Asociațiile slabe între necesitatea de încheiere cognitivă și susținerea teoriilor conspirației ar putea fi atribuite fie lipsei unei adevărate legături între cele două, fie unei asociații nonliniare. Un posibil candidat pentru o astfel de asociație nonliniară este interacțiunea dintre necesitatea de încheiere cognitivă, o predispoziție cronică de a face atribuiri conspiraționiste și dovezile disponibile în momentul în care se face o judecată. Leman și Cinnirella (2013) au abordat tocmai această problemă prin administrarea Belief in Conspiracy Theories Inventory (Inventarul încrederii în teoriile conspirației), realizat de Swami et al. (2010) ca evaluare a tendinței conspiraționiste cronice, și Need for Closure Scale unui eșantion format din studenți britanici. În plus, autorii au prezentat cazul unui accident aviatic cu cinci victime, dintre care una era un personaj politic pe care opoziția își dorea să îl blocheze. Dovezile fie implicau, fie nu implicau o conspirație.

Dacă necesitatea de finalizare cognitivă este intensă, ne putem aștepta ca audiența să dorească să ajungă la o concluzie definitivă folosind dovezile disponibile. Astfel, persoanele care au o tendință cronică de a susține teorii ale conspirației și care au și o necesitate intensă de încheiere ar trebui să prezinte o probabilitate mai mare de a pune accidentul pe seama unei conspirații, în comparație cu persoanele fără această tendință. Mai mult, participanții cu o necesitate intensă de finalizare cognitivă care nu au primit dovezi referitoare la cauza prăbușirii și la schimbări suspecte de planuri de călătorie ar trebui să presupună că la mijloc e o conspirație, în timp ce

persoanele cu o necesitate intensă de finalizare cognitivă, care au primit informația că toți investigatorii au concluzionat că prăbușirea a fost un accident, ar trebui să presupună că într-adevăr a fost vorba de un accident. Cu toate acestea, persoanele cu o necesitate scăzută de încheiere cognitivă s-ar putea să nu sară la concluzii și ar putea astfel să nu fie afectați nici de informațiile oferite, nici de propriile gânduri conspiraționiste cronice.

Rezultatele experimentului realizat de Leman și Cinnirella (2013) s-au conformat așteptărilor referitoare la tendințele conspiraționiste cronice. În general, și predispoziția constantă de a suspecta prezența conspirațiilor, și primirea de informații care susțin posibilitatea existenței unei conspirații au crescut probabilitatea ca participanții să presupună că e ceva necurat la mijloc. Cu toate acestea, necesitatea de încheiere cognitivă a moderat influența tendinței cronice de a crede în teorii ale conspirației. Persoanele cu o nevoie accentuată de încheiere cognitivă au avut o probabilitate mai mare de a face atribuiri conspiraționiste dacă aveau scoruri mari (vs. mici) la Belief in Conspiracy Theories Inventory (aproximativ $M = 34$ vs. 30 pe o scală $0-50$). Prin opoziție, persoanele cu o necesitate redusă de încheiere cognitivă au rămas indiferente la informațiile care le-au fost prezentate (aproximativ $M = 28$ în ambele cazuri). Acest studiu poate astfel sugera că necesitatea de încheiere îi face pe oameni să se bazeze pe bănuielile pe care le au în mod obișnuit pentru a explica informații ambigue, dar nu are un efect însemnat dacă informațiile sunt convingătoare.

În cadrul cercetării realizate de Marchlewska, Cichocka și Kossowska (2017), participanții polonezi au trecut prin proceduri similare celor aplicate de Leman și Cinnirella (2013). În Experimentul 1, participanții au citit informații despre venirea refugiaților în Polonia. În general, necesitatea de încheiere a corelat cu convingerea că venirea refugiaților făcea parte dintr-o conspirație împotriva Poloniei ($r=0,22$). Mai mult, efectul necesității de finalizare a fost mai mare când participanților le-a fost prezentată o conversație pe internet care sugera că venirea refugiaților făcea parte dintr-o conspirație a Uniunii Europene împotriva Poloniei. Este important că aceste efecte au fost reproduse într-un al doilea experiment, în care, în mod similar cercetării realizate de Leman și Cinnirella (2013), participanții au fost expuși la informații despre un accident aviatic.

Teoriile conspirației ca formă de susținere a eului

O motivație centrală pentru încrederea în teorii ale conspirației este, conform conceptualizării noastre, necesitatea de susținere a eului. Această motivație implică apărarea eului de gânduri neplăcute referitoare la sine, cum ar fi evitarea percepției că propria persoană este răuvoitoare, nesinceră sau incompetentă. Însă fortificarea eului implică și prezența capacităților executive, necesare pentru reușita acțiunilor și pentru gestionarea provocărilor vieții.

Apărările eului

Trăsăturile de personalitate asociate personalității schizotipale și paranoiei au dominat decenii întregi literatura asupra teoriilor conspirației. Din perspectiva teoriei psihodinamice, personalitatea schizotipală este caracterizată de omnipotență, idealizare, devalorizare, negare, proiecție primitivă sau identificare proiectivă și clivaj. Personalitatea schizotipală implică și fantasme autiste și absența refulării (Perry, Presniak și Olson, 2013). Aceste mecanisme le permit indivizilor să creadă că sinele este atotputernic și să se apere de rușine sau frică imaginându-și că aceste gânduri și emoții le aparțin celorlalți. Prin fantasme autiste, sinele poate fi protejat prin imaginarea lumii așa cum și-ar dori să fie. Când persoanele cu personalitate schizotipală pierd contactul cu realitatea, apar frecvent idei delirante de persecuție, gelozie nejustificată sau un sentiment exagerat al importanței proprii persoane. Se poate specula că schizotipia și paranoia sunt efectul acelorași tipuri de

mecanisme de apărare.

Dacă paranoia ar fi cauza teoretizării conspiraționiste, este necesar ca mai întâi să se demonstreze că cei doi factori sunt distincți. Într-o excelentă cercetare asupra acestei chestiuni, Imhoff și Lamberty (2018) au sintetizat corelațiile găsite de 11 studii între teoriile conspirației și paranoia. Aceste corelații apar în Tabelul 11.2 și indică faptul că cele două constructe sunt distincte, ceea ce Imhoff și Lamberty au confirmat prin analize factoriale în două alte studii primare.

Tabelul 11.2. Corelații între teorii ale conspirației și paranoia

Studiu	r
Barron, Morgan, Towell, Altemeyer și Swami (2014)	0,24
Brotherton și Eser (2015)	0,52
Bruder, Haffke, Neave, Nouripanah și Imhoff (2013), Studiul 2	0,45
Bruder et al. (2013), Studiul 3	0,5
Cichocka, Marchlewska și Golec de Zavala (2006)	0,37
Darwin, Neave și Holmes (2011)	0,47
Grzesiak-Feldman și Ejsmont (2008)	0,62
Grzesiak-Feldman (2015)	0,3
Wilson și Rose (2014). Studiul 1	0,27
Wilson și Rose (2014, Studiul 2	0,27
Wilson și Rose (2014, Studiul 3	0,29
Wilson și Rose (2014, Studiul 4	0,3

Sursă: Adaptat din Imhoff și Lamberty, 2018

În cadrul unuia dintre studiile incluse în meta-analiza realizată de Imhoff și Lamberty, Darwin et al. (2011) au evaluat aderența la teorii conspiraționiste printr-un chestionar alcătuit ad-hoc, care conținea itemi ca „Există servicii guvernamentale specializate care încearcă să reducă la tăcere persoanele care au văzut OZN-uri“. Autorii au folosit Paranormal Belief Scale (Scala credințelor în paranormal) (PBS; Tobacyk și Milford, 1983) pentru a evalua convingerile religioase tradiționale, credința în fenomene parapsihologice, în vrăjitorie, superstiții, spiritualism și forme de viață extraordinare și în precogniție. Au evaluat, de asemenea, tendințele paranoide prin Paranoid Ideation Scale (Scala ideăției paranoide) (Fenigstein și Vanable, 1992), care constă în 20 de afirmații cu scoruri de la 1 (deloc adevărat în cazul meu) la 5 (foarte adevărat în cazul meu), care evaluează tendința de a crede că ceilalți doresc să-i facă rău respondentului. Autorii au administrat și Schizotypal Personality Questionnaire (Chestionarul personalității schizotipale) (SPQ), care include 22 de itemi și subscale pentru deficite perceptiv-cognitive, deficite interpersonale și dezorganizare. Aceste date sunt expuse în Tabelul 11.3 și sugerează o covarianță considerabilă între teoretizarea conspiraționistă și fiecare dintre aceste scale, în special credința în fenomene parapsihologice și ideăția paranoică. Analize ale modelelor cauzale au arătat că modelul cel mai adecvat este cel în care scorul total la dimensiunea schizotipie influențează teoretizarea conspiraționistă și în mod direct, și prin medierea influențelor asupra ideăției paranoice. Însă acest model nu a inclus Paranormal Belief Scale.

Tabelul 11.3. Corelații între convingeri conspiraționiste, credințe religioase, diferite credințe în paranormal și percepții anormale

	Scor la dimen
Credințe religioase	0,26*

PSI (Parapsihologie)	0,53**
Vrăjitorie	0,40**
Spiritualism	0,43**
Superstiție	0,22*
Forme de viață extraordinare	0,30*
Precogniție	0,22**
Scor total la dimensiunea credințe în paranormal	0,47**
Ideație paranoidă	0,47**
Deficite perceptiv-cognitive	0,31**
Deficit interpersonal	0,19*
Dezorganizare	0,27*
Scor total la dimensiunea schizotipie	0,34**

*: $p < 0,05$, **: $p < 0,01$ Sursă: Adaptat din Darwin et al., 2011

Anxietatea și necesitatea de a acționa

Necesitatea de a exercita control asupra propriului mediu a fost subliniată ca factor esențial în tendința de a crede teorii ale conspirației. Conform ipotezei controlului, tendința oamenilor de a introduce scop și ordine în propria lume își are originea în necesitatea de a exercita control și facilitează încrederea în teorii conspiraționiste. Hofstadter (1964) a indicat faptul că ideile conspiraționiste ajută indivizii lipsiți de putere și de voce să înțeleagă propria realitate socială dezavantajată (vezi și Bale, 2007). Cercetări anterioare arată consecvent că motivația de a înțelege evenimentele amenințătoare din cadrul comunității determină amplificarea încrederii în teorii conspiraționiste (van Prooijen și van Dijk, 2014; vezi și van Prooijen și van Lange, 2014).

Controlul se referă la amenințări sociale, dar și la dorința generală de a menține controlul asupra oricărui aspect al lumii și a dus la cercetări asupra modalităților de creare de sens (Heine, Proulx și Vohs, 2006; Park, 2010; van den Bos, 2009) și asupra controlului compensator (Kay, Whitson, Gaucher și Galinsky, 2009; Rutjens, van Harrenveld și van der Plight, 2013). Cu cât un eveniment este mai tulburător sau cu cât este mai mare pierderea controlului, cu atât este mai probabil ca oamenii să creadă interpretări conspiraționiste ale evenimentului (Mccauley și Jacques, 1979; Pipes, 1997; Robins și Post, 1997; Shermer, 2011; J. van Prooijen și van Dijk, 2014).

Percepția controlului personal ne protejează de întâmplările aleatorii și dezorientante din viața noastră (Lerner, 1980). Potrivit lui Kay et al. (2009), oamenii au un arsenal de mecanisme psihologice compensatorii care mențin senzația de control chiar dacă de fapt controlul real nu este posibil (vezi și Axt, Landau și Kay, volumul de față). O reorganizare perceptivă compensatoare a lumii ca fiind ordonată pare o consecință a experienței anxietății, asociate reducerii controlului. Pentru a testa această posibilitate, Whitson și Galinsky (2008) le-au cerut participanților la studiu să-și amintească experiențe în care au simțit că nu au control, condiție comparată

cu situații de afirmare a sinelui. Afirmarea sinelui (Steele, 1988) este realizată adesea prin amintirea valorilor importante și este de așteptat să determine reducerea anxietății și disonanței cognitive. Astfel, în cazul activării sentimentelor asociate reducerii controlului, afirmarea sinelui ar trebui să restaureze senzația de control. În acest studiu, participanții care nu dețineau resurse pentru afirmarea sinelui au avut o tendință mai accentuată de a crede explicații conspiraționiste decât cei cărora li s-a cerut să-și amintească alte tipuri de materiale sau nimic. Cu alte cuvinte, restabilirea încrederii în sine a redus anxietatea și a eliminat necesitatea de a vedea ordine în evenimente ambigue.

O limită a datelor din studiul realizat de Whitson și Galinsky este absența unui nivel de referință pentru manipularea controlului. Însă, din fericire, van Prooijen și Acker (2015) au realizat un experiment similar în cadrul căruia participanții au scris despre o situație în care au simțit că aveau control, că nu aveau control sau în care nu era vorba despre control. Apoi, participanții au citit informații despre construirea unei linii de metrou în Amsterdam și le-au fost puse întrebări despre corupția din protocoalele de acordare a contractului de construcție. La fel ca în cercetarea realizată de Whitson și Galinsky, participanții cu un nivel redus de control au avut o tendință mai accentuată de a crede în conspirații decât participanții cu un nivel ridicat de control. Descoperirilor anterioare li s-a adăugat informația că participanții din grupul de control nu au prezentat diferențe semnificative în comparație cu participanții cu nivel scăzut de control, dar au prezentat diferențe în comparație cu participanții cu nivel ridicat de control. În concluzie, amenințările asupra controlului amplifică tendința de a da lumii scop, sens și ordine, cu prețul credinței în conspirații.

Este interesant că van Prooijen și Acker (2015) au reanalizat și o cercetare asupra conspirațiilor legate de Y2K*****. Acest studiu a inclus evaluări ale amenințării pe care eroarea de software Y2K o reprezenta pentru funcționarea sistemelor informatice din întreaga lume, precum și diferite credințe conspiraționiste, printre acestea fiind asasinarea lui Kennedy și ascunderea dovezilor că există extratereștri. Gradul în care respondenții se simțeau amenințați de Y2K corela cu credința lor în diferite teorii ale conspirației care nu aveau nicio legătură cu Y2K, ceea ce sugerează că este probabil ca orice amenințare percepută să provoace explicații conspiraționiste

ale evenimentelor. Este interesant că amenințarea percepută ca prezentată de Y2K corela negativ cu credința că această eroare de software era în centrul unei conspirații. Această asociație negativă poate sugera că oamenii care se simțeau în mai mare măsură amenințați erau mai bine informați despre cauzele Y2K. Altfel, legătura evidentă dintre cele două variabile i-ar fi putut face pe participanți să-și schimbe raționamentul referitor la conspirație, ca efort de a corecta influența propriilor emoții.

O demonstrație mai degrabă generală a asociațiilor dintre încrederea într-o teorie a conspirației și stres a fost oferită într-o cercetare de corelație realizată de Swami și colegii săi (2016). Cercetătorii au colectat date de la adulți americani înscriși la Amazon Mechanical Turk. Participanții au completat Belief in Conspiracy Theories Inventory, precum și Perceived Stress Scale (Scala stresului perceput) (Cohen, Kamarck și Mermelstein, 1983), List of Threatening Experiences Questionnaire (Chestionarul experiențelor amenințătoare) (Brugha, Bebbington, Tennant și Hurry, 1985), versiunea Y-1 a State-Trait Anxiety Inventory (Inventarul de anxietate stare-trăsătură) (Spielberger, Gorsuch, Lushene, Vagg și Jacobs, 1983) ca evaluare a anxietății ca stare, versiunea Y-2 a State-Trait Anxiety Inventory (Spielberger et al., 1983) ca evaluare a anxietății ca trăsătură și Profile of Mood States (Profilul stărilor dispoziționale) (Shacham, 1983). Stresul perceput, evenimentele de viață stresante și anxietatea ca trăsătură au corelat pozitiv cu încrederea în teorii ale conspirației (valorile r s-au situat între 0,10 și 0,29).

Una dintre formele anxietății care a primit atenție în literatură a fost atașamentul nesecurizant. Persoanele cu atașament securizant (vs. nesecurizant) caută susținere instrumentală și afectivă din partea celorlalți (Florian, Mikulincer și Taubman, 1995; Larose, Bernier, Soucy și Duchesne, 1999). Persoanele cu atașament nesecurizant tind să evite (atașament evitant) sau caută intens apropierea celorlalți. Indivizii anxioși tind să exagereze amenințările ca formă de a obține susținere din partea celorlalți și sunt hipervigilenți în domeniul interpersonal (Cassidy și Berlin, 1994; Cassidy și Kobak, 1988; Mikulincer, Shaver și Pereg, 2003). Green și Douglas (2018) au investigat posibilitatea ca hipervigilența asociată atașamentului anxios să fie asociată cu încrederea în teorii conspiraționiste ca încercare de reducere a anxietății (Douglas, Sutton și Cichocka, 2017). Pentru a testa această ipoteză, Green și Douglas au administrat chestionare care evaluează atașamentul și

credința în teorii ale conspirației și în tendința generală a grupurilor care dețin puterea de a conspira. Conform așteptărilor, atașamentul anxios a prezis gândirea conspiraționistă chiar și după controlul altor variabile interpersonale și politice.

Efectele determinate de nivelul redus de control, de anxietate, de stres și de stilul de atașament asupra tendinței de a crede teorii ale conspirației au o probabilitate ridicată nu doar de a fi neplăcute, ci și de a fi problematice pentru acțiunile propriu-zise. De fapt, s-a propus ipoteza conform căreia disonanța cognitivă provoacă efecte nu din cauza experienței anxietății, ci din cauza perturbării acțiunii cauzate de aceste sentimente (Harmon-Jones, Harmon-Jones și Levy, 2015). Întrucât convingerile și atitudinile determină automat acțiuni, apar conflicte dacă aceste convingeri și atitudini au implicații contradictorii. Astfel, susținerea acțiunilor într-o modalitate relativ ordonată este probabil să reprezinte un factor care stă la baza multor descoperiri asociate reducerii controlului perceput.

Stima de sine și narcisismul

S-a susținut că menținerea stimei de sine stă la baza teoriilor conspirației, ipoteza fiind că explicațiile și plasarea aspectelor negative în afara sinelui sunt reconfortante (Robins și Post, 1997). Mai multe studii au evaluat stima de sine și au corelat-o cu evaluări ale teoriilor conspirației. De exemplu, în cadrul unui studiu realizat în Marea Britanie și Austria (Swami et al., 2011), participanții au completat instrumente de evaluare referitoare la teoria conspirației și stima de sine. Credința într-o conspirație în jurul atacurilor cu bombe din Londra de la 7 iulie 2005 și un inventar general de încredere în teoriile conspirației au corelat cu stima de sine ($r = -0,16$ și $r = -0,20$), ceea ce susține ipoteza că persoanele cu stimă de sine scăzută susțin ideile conspiraționiste în mai mare măsură decât persoanele cu stimă de sine ridicată, ceea ce s-a reprodus într-un al doilea studiu care a folosit o teorie fictivă a conspirației, precum și în alte cercetări (Galliford și Furnham, 2017).

Această descoperire aparent robustă are însă două explicații posibile. Pe de o parte, stima de sine redusă poate amplifica aceste credințe ca încercare de a accentua sentimentele pozitive față de sine. Pe de altă parte, stima de sine scăzută poate pur și simplu să coloreze perspectiva oamenilor asupra lumii într-un mod negativ.

Cercetarea realizată de Cichocka, Marchlewska, De și Olechowski (2016) a pus în lumină procesele prin care stima de sine influențează ideile conspiraționiste. Chiar dacă corelațiile de ordin zero între stima de sine și teoriile conspirației nu sunt semnificative din punct de vedere statistic, stima de sine prezice credința în teoriile conspirației dacă este controlat narcisismul. Însă este interesant că narcisismul are o asociație pozitivă consecventă cu teoriile conspirației. Specific, narcisicii au o probabilitate mai mare decât non-narcisicii de a se considera invidiați și obiecte ale conspirațiilor din partea celorlalți, idee paranoidă care prezice credința în teoriile conspirației.

Cele trei studii raportate de Cichocka et al. (2016) susțin ideea că narcisismul corelează pozitiv cu teoriile conspirației (e.g. $r = 0,25$ în Studiul 2) și paranoia (e.g. $r = 0,18$ în Studiul 2) și că unele influențe ale narcisismului asupra teoriilor conspirației sunt mediate de paranoia. Prin opoziție, stima scăzută de sine este asociată cu niveluri mai mici de paranoia, posibil pentru că e nevoie ca persoana să se considere valoroasă ca să se considere în centrul unei conspirații. Mai mult, stima de sine a corelat cu respectul pentru omenire, ceea ce implică faptul că autovalorizarea prezice o atitudine generală pozitivă față de ceilalți.

Toate cercetările asupra stimei de sine și narcisismului sunt firește de natură corelativă, dar variabila inherentă a necesității de unicitate a fost investigată prin proceduri experimentale. Necesitatea de a fi unic este dorința de a fi diferit de ceilalți și de a ieși în evidență prin caracteristici și idei unice (Lynn și Snyder, 2002). Teoriile conspirației sunt indicatori ideali ai unicității, dat fiind că sunt relativ neconvenționale și sugerează acces la o cunoaștere privilegiată (Lantian, Muller, Nurra și Douglas, 2017). Participanții care au scris despre o situație anterioară în care și-au exprimat individualitatea, prin opoziție cu cei care au scris despre exprimarea conformismului, au raportat interpretări mai accentuat conspiraționiste ale unui accident (Lantian et al.,

2017). În această lumină, pare posibil ca asociația pozitivă dintre narcisism și teoriile conspirației să fie determinată și de numeroasele tendințe ale narcisicilor de a se considera superiori din punct de vedere intelectual, și astfel mai capabili să perceapă intrigi dificil de descoperit (pentru cercetări în domeniul narcisismului colectiv, vezi Golec de Zavala și Federico, 2018).

Teoriile conspirației ca formă de integrare socială

Teoriile conspirației sunt un fenomen social, în timp ce paranoia se referă la idei individuale. Astfel, la fel ca în cazul altor fenomene sociale, trebuie să luăm în considerare motivele sociale care susțin gândirea de tip conspiraționist, multe dintre acestea fiind investigate în literatură. Desigur, numeroși factori sociali implicați în teorii ale conspirației sunt de natură politică, dar efectele izolării și excluderii au legătură în general cu necesitatea de integrare socială în rețele informale.

Izolarea și excluderea socială au reprezentat principalele variabile explicative în studierea teoriilor conspirației, dar dovezile nu sunt lipsite de ambiguități. Pe de o parte, izolarea și excluderea socială par să susțină credința în teorii ale conspirației prin amplificarea anxietății și necesității de a simplifica lumea. Graeupner și Coman (2017) au investigat această posibilitate prin două studii, unul de corelații și unul experimental. Corelația dintre sentimentele de excludere și susținerea teoriilor conspirației a fost $r = 0,19$. Pe de altă parte, legăturile sociale și includerea în grupuri pot accentua preocuparea pentru amenințări sociale cum sunt cele prezentate de teoriile conspirației. De fapt, al doilea studiu realizat de Graeupner și Coman a arătat că alegerea evidentă a unei persoane ca partener într-o sarcină experimentală (manipularea excluderii) a dus la atribuiri conspiraționiste mai accentuate decât faptul de a nu fi selectat, efect egal cu $r = 0,25$.

Rezumat

Trecerea în revistă din acest capitol sugerează că încep să fie din ce în ce mai multe cercetări asupra modului în care argumentele devin plauzibile și nefalsificabile și că este necesar ca oamenii de știință să abordeze cu atenție aceste subiecte. În speță, numeroase studii din literatură sunt realizate cu studenți și eșantioane nereprezentative. Mai mult, cercetările nu au luat suficient în considerare gradul în care social media și alte tehnologii digitale contribuie la diseminarea teoriilor conspirației și a altor forme periculoase de fake news. În viitor, eșantioane naționale mai diverse și studiul diferitelor tehnologii va face mai multă lumină în epoca post-adevăr în care trăim.

În acest capitol am sugerat că teoriile conspirației reprezintă un caz fascinant de argument susținut și pe baza verificabilității, și a neverificabilității. Acești doi poli fac legătura între teoriile conspirației și necesitățile umane absolut fundamentale: necesitatea de cunoaștere, necesitatea de apărare a eului și necesitatea de integrare socială. Astfel, teoriile conspirației oferă o pseudo-realitate plauzibilă sau aproape științifică și, în același timp, mențin misterul mușamalizărilor. Această asociere face astfel de convingeri durabile și dificil de corectat.

Referințe

Bale, J.M. (2007). Political paranoia v. political realism: On distinguishing between bogus conspiracy theories and genuine conspiratorial politics: Patterns of prejudice (Vol. 41). Graduate School of International Policy Studies, Monterey Institute of International Studies. Monterey, CA, US: Taylor & Francis. <http://dx.doi.org/10.1080/00313220601118751>

Barron, D., Morgan, K., Towell, T., Altemeyer, B. și Swami, V. (2014).

Associations between schizotypy and belief in conspiracist ideation: Personality and individual differences (Vol. 70). Department of Psychology, University of Westminster. Londra, Marea Britanie
david.barron@my.westminster.ac.uk; Barron, David, 309 Regent Street, Londra, Marea Britanie, W1B 2UW: Department of Psychology, University of Westminster, david.barron@my.westminster.ac: Elsevier Science.
<http://dx.doi.org/10.1016/j.paid.2014.06.040>

Baumeister, R.F. (1997). Identity, self-concept, and self-esteem: The self lost and found. În Handbook of personality psychology (pp. 681–710). San Diego, CA, USA: Academic Press.

Braswell, G.W. (2000). What you need to know about Islam & Muslims. Nashville, TN: B&H Publishing Group.

Brotherton, R. și Eser, S. (2015). Bored to fears: Boredom proneness, paranoia, and conspiracy theories. Personality and Individual Differences, 80, 1–5. <https://doi.org/10.1016/j.paid.2015.02.011>

Bruder, M., Haffke, P., Neave, N., Nouripanah, N. și Imhoff, R. (aprilie 2013). Measuring individual differences in generic beliefs in conspiracy theories across cultures: Conspiracy mentality questionnaire. Frontiers in Psychology, 4, 225. <https://doi.org/10.3389/fpsyg.2013.00225>

Brugha, T., Bebbington, P., Tennant, C. și Hurry, J. (1985). The list of threatening experiences: A subset of 12 life event categories with considerable long-term contextual threat. Psychological Medicine. <https://doi.org/10.1017/S003329170002105X>

Buhr, K. și Dugas, M.J. (2002). The intolerance of uncertainty scale: Psychometric properties of the English version. Behaviour Research and Therapy, 40 (8), 931–945. [https://doi.org/10.1016/S0005-7967\(01\)00092-4](https://doi.org/10.1016/S0005-7967(01)00092-4)

Business Insider. (2018). 19 outlandish conspiracy theories Donald Trump has floated on the campaign trail and in the White House. Accesat la www.businessinsider.com/donald-trump-conspiracy-theories-2016-5#claims-3000-people-didnt-die-inpuerto-rico-after-hurricane-maria-and-that-democrats-inflated-the-death-toll-19

Carleton, R.N., Norton, M.A.P.J. și Asmundson, G.J.G. (2007). Fearing the unknown: A short version of the Intolerance of Uncertainty Scale. *Journal of Anxiety Disorders*. <https://doi.org/10.1016/j.janxdis.2006.03.014>

Cassidy, J. și Berlin, L.J. (1994). The insecure/ambivalent pattern of attachment: Theory and research. *Child Development*. <https://doi.org/10.1111/j.1467-8624.1994.tb00796.x>

Cassidy, J. și Kobak, R.R. (1988). Avoidance and its relation to other defensive process. În J. Belsky și T. Nezworski (coord.), *Child psychology. Clinical implications of attachment* (pp. 300–323). Mahwah, NJ: Lawrence Erlbaum Associates.

Chaiken, S. și Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/0022-3514.66.3.460>

Chan, M.S., Jones, C.R. și Albarracín, D. (2017). Countering false beliefs: An analysis of the evidence and recommendations of best practices for the retraction and correction of scientific misinformation. În *Handbook of the science of science communication*. <https://doi.org/10.1093/oxfordhb/9780190497620.013.37>

Chen, S. și Chaiken, S. (1999). The heuristic-systematic model in its broader context. În *Dualprocess theories in social psychology*. <https://doi.org/10.4319/lo.2013.58.2.0489>

Chen, S., Duckworth, K. și Chaiken, S.R. (1999). Motivated heuristic and systematic processing. *Psychological Inquiry*. https://doi.org/10.1207/s15327965pli1001_6

Cichocka, A., Marchlewska, M., De, A.G. și Olechowski, M. (2016). „They will not control us”: Ingroup positivity and belief in intergroup conspiracies, 556–576. <https://doi.org/10.1111/bjop.12158>

Cichocka, A., Marchlewska, M. și Golec de Zavala, A. (2016). Does self-love or self-hate predict conspiracy beliefs? Narcissism, self-esteem, and the

endorsement of conspiracy theories: Social psychological and personality science (Vol. 7). Canterbury, UK: School of Psychology, University of Kent a.k.cichocka@kent.ac.uk; Institute for Social Studies, University of Warsaw, Warsaw, Poland; Department of Psychology, Goldsmiths, University of London, Londra, United Kingdom; Cichock: Sage Publications.
<http://dx.doi.org/10.1177/1948550615616170>

Clarke, S. (2002). Conspiracy theories and conspiracy theorizing. *Philosophy of the Social Sciences*. <https://doi.org/10.1177/004931032002001>

Cohen, S., Kamarck, T. și Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24 (4), 385–396.

Darwin, H., Neave, N. și Holmes, J. (2011). Belief in conspiracy theories. The role of paranormal belief, paranoid ideation and schizotypy. *Personality and Individual Differences*, 50 (8), 1289–1293.
<https://doi.org/10.1016/j.paid.2011.02.027>

Davies W. (24 august 2016). The age of post-truth politics. *The New York Times*. Accesat la www.nytimes.com/

Delong-Bas, N.J. (2008). *Wahhabi Islam: From revival and reform to global jihad*. Oxford: Oxford University Press.

Douglas, K.M., Sutton, R.M. și Cichocka, A. (2017). The psychology of conspiracy theories. *Current Directions in Psychological Science*, 26 (6), 538–542. <https://doi.org/10.1177/0963721417718261>

Douglas, K.M., Sutton, R.M., Jolley, D. și Wood, M.J. (2015). The social, political, environmental, and health-related consequences of conspiracy theories. În M. Bilewicz, A. Cichocka și W. Soral (coord.), *The psychology of conspiracy* (pp. 183–200). New York și Londra: Routledge/Taylor & Francis Group.

Eagly, A.H. și Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich College Publishers.

Eggertson, L. (2010). Lancet retracts 12-year-old article linking autism to

MMR vaccines. CMAJ: Canadian Medical Association Journal = Journal de l'Association Médicale Canadienne, 182 (4).
<https://doi.org/10.1503/cmaj.109-3179>

Enster, M. (1999). Conspiracy theories: Secrecy and power in American culture. St. Paul, MN: University of Minnesota Press.

Fenigstein, A. și Vanable, P.A. (1992). Paranoia and self-consciousness. Journal of Personality and Social Psychology. <https://doi.org/10.1037/0022-3514.62.1.129>

Florian, V., Mikulincer, M. și Taubman, O. (1995). Does hardiness contribute to mental health during a stressful real-life situation? The roles of appraisal and coping. Journal of Personality and Social Psychology.
<https://doi.org/10.1037/0022-3514.68.4.687>

Freeston, M.H., Rhéaume, J., Letarte, H., Dugas, M.J. și Ladouceur, R. (1994). Why do people worry? Personality and Individual Differences.
[https://doi.org/10.1016/0191-8869\(94\)90048-5](https://doi.org/10.1016/0191-8869(94)90048-5)

Galliford, N. și Furnham, A. (2017). Individual difference factors and beliefs in medical and political conspiracy theories: Scandinavian journal of psychology (Vol. 58). Research Department of Clinical, Educational and Health Psychology, University College London. Londra, Marea Britanie: Wiley-Blackwell Publishing Ltd. <http://dx.doi.org/10.1111/sjop.12382>

Golec de Zavala, A. și Federico, C.M. (2018). Collective narcissism and the growth of conspiracy thinking over the course of the 2016 United States presidential election: A longitudinal analysis. European Journal of Social Psychology. <https://doi.org/10.1002/ejsp.2496>

Graeupner, D. și Coman, A. (2017). The dark side of meaning-making: How social exclusion leads to superstitious thinking: Journal of experimental social psychology (Vol. 69). Princeton, NJ, US: Department of Psychology, Princeton University acoman@princeton.edu; Coman, Alin, Peretsman-Scully Hall, #529, Princeton, US, 08540, Department of Psychology, Princeton University, acoman@princeton.edu: Elsevier Science.
<http://dx.doi.org/10.1016/j.jesp.2016.10.003>

Green, R. și Douglas, K.M. (2018). Anxious attachment and belief in conspiracy theories. *Personality and Individual Differences*, 125, 30–37. <https://doi.org/10.1016/j.paid.2017.12.023>

Groh, D. (1987). Draft of a theory of conspiracy theories. În *Changing conceptions of conspiracy*. <https://doi.org/10.1007/978-1-4612-4618-3>

Grzesiak-Feldman, M. (2015). Are the high authoritarians more prone to adopt conspiracy theories? The role of right-wing authoritarianism in conspiratorial thinking. (M. Bilewicz, A. Cichocka, și W. Soral, coord.), *The psychology of conspiracy BT: The psychology of conspiracy*. New York, NY: Routledge and Taylor & Francis Group. Accesat la <https://search.proquest.com/docview/1717508783?accountid=14553>

Grzesiak-Feldman, M. și Ejsmont, A. (2008). Paranoia and conspiracy thinking of Jews, Arabs, Germans, and Russians in a Polish sample: *Psychological reports* (Vol. 102). Warsaw, Poland: University of Warsaw ika@psych.uw.edu.pl; Grzesiak-Feldman, Monika, Stawki 5/7, Warsaw, Poland, 00–183, Faculty of Psychology, University of Warsaw, ika@psych.uw.edu.pl: *Psychological Reports* Sage Publications. <http://dx.doi.org/10.2466/PRO.102.3.884-886>

Harmon-Jones, E., Harmon-Jones, C. și Levy, N. (2015). An action-based model of cognitive-dissonance processes. *Current Directions in Psychological Science*. <https://doi.org/10.1177/0963721414566449>

Hart, W., Albarracín, D., Eagly, A.H., Lindberg, M., Lee, K.H. și Brechan, I. (2009). Feeling validated vs. being correct: A meta-analysis of exposure to information. *Psychological Bulletin*, 135, 555–588. doi: 10.1037/a0015701. PMID: 19586162

Heine, S.J., Proulx, T. și Vohs, K.D. (2006). The meaning maintenance model: On the coherence of social motivations. *Personality and Social Psychology Review*, 10 (2), 88–110. https://doi.org/10.1207/s15327957pspr1002_1

Heller, J. (2015). Rumors and realities: Making sense of HIV/AIDS conspiracy narratives and contemporary legends. *American Journal of Public*

Health, 105 (1), e43–e50. <https://doi.org/10.2105/AJPH.2014.302284>

Hofstadter, R. (noiembrie 1964). The paranoid style in American politics. Harpers Magazine.

Huff, T.E. (1995). Rethinking islam and fundamentalism. Sociological Forum. <https://doi.org/10.1007/BF02095834>

Imhoff, R. și Lamberty, P. (2018). How paranoid are conspiracy believers? Toward a more fine-grained understanding of the connect and disconnect between paranoia and belief in conspiracy theories. European Journal of Social Psychology. Mainz, Germany: Johannes Gutenberg University of Mainz roland.imhoff@uni-mainz.de: John Wiley & Sons. <http://dx.doi.org/10.1002/ejsp.2494>

Janoff-Bulman, R. (1989). Assumptive worlds and the stress of traumatic events: Applications of the schema construct. Social Cognition. <https://doi.org/10.1521/soco.1989.7.2.113>

Kay, A.C., Whitson, J.A., Gaucher, D. și Galinsky, A.D. (2009). Compensatory control: Achieving order through the mind, our institutions, and the heavens: Current directions in psychological science (Vol. 18). Waterloo, ON, Canada: University of Waterloo ackay@uwaterloo.ca; University of Texas at Austin, Austin, TX, US; Northwestern University, Evanston, IL, US; Kay, Aaron C., 200 University AveWest, Waterloo, Canada, Department of Psychology, University of Water: Wiley-Blackwell Publishing Ltd. Blackwell Publishing Sage Publications. <http://dx.doi.org/10.1111/j.1467-8721.2009.01649.x>

Kruglanski, A.W. (1989). Lay epistemics and human knowledge: Cognitive and motivational bases: Perspectives in social psychology. <https://doi.org/10.1007/978-1-4899-0924-4>

Kruglanski, A.W. (1990). Lay epistemic theory in social-cognitive psychology. Psychological Inquiry. https://doi.org/10.1207/s15327965pli0103_1

Kruglanski, A.W., Dechesne, M., Orehek, E. și Pierro, A. (2009). Three

decades of lay epistemics: The why, how, and who of knowledge formation. *European Review of Social Psychology*.
<https://doi.org/10.1080/10463280902860037>

Kruglanski, A.W., Orehek, E., Dechesne, M. și Pierro, A. (2010). Lay epistemic theory: The motivational, cognitive, and social aspects of knowledge formation. *Social and Personality Psychology Compass*.
<https://doi.org/10.1111/j.1751-9004.2010.00308.x>

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*.
<https://doi.org/10.1037/0033-2909.108.3.480>

Lantian, A., Muller, D., Nurra, C. și Douglas, K.M. (2017). „I know things they don't know!": The role of need for uniqueness in belief in conspiracy theories: *Social psychology* (Vol. 48). Grenoble, France: University Grenoble Alpes dominique.muller@univgrenoble-alpes.fr; School of Psychology, University of Kent, United Kingdom; Muller, Dominique, Bâtiment BSHM, Grenoble Cedex 9, France, 38 058, LIP/PC2S, University Grenoble Alpes, dominique.muller@univ-: Hogrefe Publishing Hogrefe & Huber Publishers Verlag Hans Huber. <http://dx.doi.org/10.1027/1864-9335/a000306>

Larose, S., Bernier, A., Soucy, N. și Duchesne, S. (1999). Attachment style dimensions, network orientation and the process of seeking help from college teachers. *Journal of Social and Personal Relationships*.
<https://doi.org/10.1177/0265407599162006>

Leman, P.J. și Cinnirella, M. (2013). Beliefs in conspiracy theories and the need for cognitive closure. *Frontiers in Psychology*, 4.
<https://doi.org/10.3389/fpsyg.2013.00378>

Lerner, M.J. (1980). The belief in a just world. În *The belief in a just world: A fundamental delusion*. https://doi.org/10.1007/978-1-4899-0448-5_2

Lynn, M. și Snyder, C.R. (2002). Uniqueness seeking. În C.R. Snyder și S.J. Lopez (coord.), *Handbook of positive psychology* (pp. 395–410). New York, NY: Oxford University Press.

Marchlewska, M., Cichocka, A. și Kossowska, M. (2017). Addicted to

answers: Need for cognitive closure and the endorsement of conspiracy beliefs. *European journal of social psychology* (Vol. 48). Warsaw, Poland: Institute for Social Studies, University of Warsaw
marta.marchlewska@psych.uw.edu.pl; School of Psychology, University of Kent, Canterbury, United Kingdom; Institute of Psychology, Jagiellonian University, Kraków, Poland: John Wiley & Sons.
<http://dx.doi.org/10.1002/ejsp.2308>

Mccauley, C. și Jacques, S. (1979). The popularity of conspiracy theories of presidential assassination: A Bayesian analysis. *Journal of Personality and Social Psychology*, 37 (5), 637–644. <https://doi.org/10.1037/0022-3514.37.5.637>

Meagher, P.K., și O'Brien, T.C. (1979). *Encyclopedic dictionary of religion* (Vol. 3). Sisters of St. Joseph of Philadelphia. Washington, DC: Corpus Publications.

Mikulincer, M., Shaver, P.R. și Pereg, D. (2003). Attachment theory and affect regulation: The dynamics, development, and cognitive consequences of attachment-related strategies. *Motivation and Emotion*.
<https://doi.org/10.1023/A:1024515519160>

Miller, S. (2002). Conspiracy theories: Public arguments as coded social critiques: A Rhetorical analysis of the TWA flight 800 conspiracy theories. *Argumentation and Advocacy*.
<https://doi.org/10.1080/00028533.2002.11821576>

Moulding, R., Nix-Carnell, S., Schnabel, A., Nedeljkovic, M., Burnside, E.E., Lentini, A.F. și Mehzabin, N. (2016). Better the devil you know than a world you don't? Intolerance of uncertainty and worldview explanations for belief in conspiracy theories: Personality and individual differences (Vol. 98). Burwood, VIC, Australia: School of Psychology, Deakin University
richard.moulding@deakin.edu.au; School of Psychology, Swinburne University of Technology, VIC, Australia; Moulding, Richard, Melbourne Burwood Campus, 221 Burwood Highway, Burwood, Australia, 3: Elsevier Science. <http://dx.doi.org/10.1016/j.paid.2016.04.060>

Nasir, J.J. (2009). The status of women under Islamic law and modern

Islamic legislation. Boston, MA: Brill.

Palmer, R. și Albarracín, D. (2018). Trust in science as a deterrent and a facilitator of belief in conspiracy theories: Pseudoscience preys on audiences that trust in science. University of Illinois at Urbana-Champaign. https://doi.org/10.13012/B2IDB-4469040_V1

Park, C.L. (2010). Making sense of the meaning literature: An integrative review of meaning making and its effects on adjustment to stressful life events. *Psychological Bulletin*. <https://doi.org/10.1037/a0018301>

Perry, J.C., Presniak, M.D. și Olson, T.R. (2013). Defense mechanisms in schizotypal, borderline, antisocial, and narcissistic personality disorders. *Psychiatry: Interpersonal and Biological Processes*, 76 (1), 32–52. <https://doi.org/10.1521/psyc.2013.76.1.32>

Pipes, D. (1997). *Conspiracy: How the paranoid style flourishes and where it comes from*. New York: Free Press.

Politico. (2011). Birthersim: Where it all began. Politico. În B. Smith, și B. Tau. Accesat la <https://www.politico.com/story/2011/04/birtherism-where-it-all-began-053563>

Pratt, R. (2003). Theorizing conspiracy. *Theory and Society*. <https://doi.org/10.1023/A:1023996501425>

Prooijen, J.-W. și Acker, M. (2015). The influence of control on belief in conspiracy theories: Conceptual and applied extensions: *Applied cognitive psychology* (Vol. 29). Amsterdam, Netherlands: Netherlands Institute for the Study of Crime and Law Enforcement j.w.van.prooijen@vu.nl; Otterbein University, Westerville, OH, US; Prooijen, Jan-Willem, Van der Boechorststraat 1, Amsterdam, Netherlands, 1081 BT, Department of Social: John Wiley & Sons. <http://dx.doi.org/10.1002/acp.3161>

Robins, R.S. și Post, J.M. (1997). *Political paranoia: The psychopolitics of hatred*. New Haven, CT: Yale University Press.

Roets, A. și Van Hiel, A. (2011). Item selection and validation of a brief, 15-

item version of the need for closure scale. *Personality and Individual Differences*. <https://doi.org/10.1016/j.paid.2010.09.004>

Rolling Stones. (2017). Anatomy-of-a-fake-news-scandal. Accesat la www.rollingstone.com/politics/politics-news/anatomy-of-a-fake-news-scandal-125877/

Rutjens, B.T., van Harreveld, F. și van der Pligt, J. (2013). Step by step: Finding compensatory order in science. *Current Directions in Psychological Science*. <https://doi.org/10.1177/0963721412469810>

Shacham, S. (1983). A shortened version of the profile of mood states. *Journal of Personality Assessment*. https://doi.org/10.1207/s15327752jpa4703_14

Shermer, M. (2011). *The believing brain: From ghosts and gods to politics and conspiracies: How we construct beliefs and reinforce them as truths*. New York, NY, USA: Times Books.

Sorrentino, R.M. și Short, J.A.C. (1986). Uncertainty orientation, motivation, and cognition. În R.M. Sorrentino și E.T. Higgins (coord.), *The handbook of motivation and cognition: Foundations of social behavior* (Vol. 1; pp. 379–403). New York: Guilford Press.

Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R. și Jacobs, G.A. (1983). *Manual for the State-Trait Anxiety Inventory (STAI Form Y)*: Consulting psychologists Palo Alto. <https://doi.org/10.1002/9780470479216.corpsy0943>

Steele, C.M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. *Advances in Experimental Social Psychology*. [https://doi.org/10.1016/S0065-2601\(08\)60229-4](https://doi.org/10.1016/S0065-2601(08)60229-4)

Swami, V., Chamorro-Premuzic, T. și Furnham, A. (2010). Unanswered questions: A preliminary investigation of personality and individual difference predictors of 9/11 conspiracist beliefs: *Applied cognitive psychology* (Vol. 24). London, UK: Department of Psychology, University of Westminster virenswami@hotmail.com; Department of Psychology,

Goldsmiths, University of London, London, United Kingdom; Department of Psychology, University College London, London, United Kingdom: John Wiley & Sons. <http://dx.doi.org/10.1002/acp.1583>

Swami, V., Coles, R., Stieger, S., Pietschnig, J., Furnham, A., Rehim, S. și Voracek, M. (2011). Conspiracist ideation in Britain and Austria: Evidence of a monological belief system and associations between individual psychological differences and realworld and fictitious conspiracy theories. *British Journal of Psychology*, 102 (3), 443–463. <https://doi.org/10.1111/j.2044-8295.2010.02004.x>

Swami, V., Furnham, A., Smyth, N., Weis, L., Lay, A. și Clow, A. (2016). Putting the stress on conspiracy theories: Examining associations between psychological stress, anxiety, and belief in conspiracy theories: *Personality and individual differences* (Vol. 99). Cambridge, UK: Department of Psychology, Anglia Ruskin University viren.swami@anglia.ac.uk; Department of Clinical, Educational, and Health Psychology, University College London, London, United Kingdom; Department of Psychology, University of: Elsevier Science. <http://dx.doi.org/10.1016/j.paid.2016.04.084>

Tesser, A. (2001). On the plasticity of self-defense. *Current Directions in Psychological Science*. <https://doi.org/10.1111/1467-8721.00117>

Tobacyk, J. și Milford, G. (1983). Belief in paranormal phenomena: Assessment instrument development and implications for personality functioning. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/0022-3514.44.5.1029>

van den Bos, K. (2009). Making sense of life: The existential self trying to deal with personal uncertainty. *Psychological Inquiry*. <https://doi.org/10.1080/10478400903333411>

Van der Wal, R., Sutton, R.M. Lange, J. și Braga, J. (2018). Suspicious binds: Conspiracy thinking and tenuous perceptions of causal connections between cooccurring and spuriously correlated events. *European Journal of Social Psychology*, 48, 970–989.

van Harreveld, F., Rutjens, B.T., Schneider, I.K., Nohlen, H.U. și Keskinis,

K. (2014). In doubt and disorderly: Ambivalence promotes compensatory perceptions of order: *Journal of experimental psychology: General* (Vol. 143). Amsterdam, Netherlands: Department of Social Psychology, University of Amsterdam f.vanharreveld@uva.nl; Department of Clinical Psychology, VU University of Amsterdam, Amsterdam, Netherlands; van Harreveld, Frenk, Weesperplein 4, Amsterdam, Netherlands, 1018: American Psychological Association Psychological Review Company.
<http://dx.doi.org/10.1037/a0036099>

van Prooijen, J.-W., Douglas, K. și De Inocencio, C. (2018). Connecting the dots: Illusory pattern perception predicts beliefs in conspiracies and the supernatural. *European Journal of Social Psychology*, 48, 320–335.

van Prooijen, J.-W. și Jostmann, N.B. (2013). Belief in conspiracy theories: The influence of uncertainty and perceived morality: *European journal of social psychology* (Vol. 43). Amsterdam, Netherlands: Department of Social and Organizational Psychology, VU University Amsterdam
jw.van.prooijen@psy.vu.nl; University of Amsterdam, Amsterdam, Netherlands; van Prooijen, Jan-Willem, Van den Boechorststraat 1, Amsterdam, Netherlands, 1081: John Wiley & Sons.
<http://dx.doi.org/10.1002/ejsp.1922>

van Prooijen, J.-W. și Krouwel, A.P.M. (2015). Mutual suspicion at the political extremes: How ideology predicts belief in conspiracy theories. În M. Bilewicz, A. Cichocka și W. Soral (coord.), *The psychology of conspiracy* (pp. 79–98). New York și London: Routledge/Taylor & Francis Group.

van Prooijen, J.-W. și van Dijk, E. (2014). When consequence size predicts belief in conspiracy theories: The moderating role of perspective taking. *Journal of Experimental Social Psychology*, 55, 63–73.
<https://doi.org/10.1016/j.jesp.2014.06.006>

van Prooijen, J.-W. și van Lange, P.A.M. (2014). The social dimension of belief in conspiracy theories. În *Power, politics, and paranoia: Why people are suspicious of their leaders*.
<https://doi.org/10.1017/CBO9781139565417.017>

Webster, D.M. și Kruglanski, A.W. (1994). Individual differences in need for

cognitive closure. *Journal of Personality and Social Psychology*.
<https://doi.org/10.1037/0022-3514.67.6.1049>

Whitson, J.A. și Galinsky, A.D. (2008). Lacking control increases illusory pattern perception. *Science*. <https://doi.org/10.1126/science.1159845>

Wichman, A.L., Brinol, P., Petty, R.E., Rucker, D.D., Tormala, Z.L. și Weary, G. (2010). Doubting one's doubt: A formula for confidence? *Journal of Experimental Social Psychology*, 46 (2), 350–355.
<https://doi.org/10.1016/j.jesp.2009.10.012>

Wilson, M.S. și Rose, C. (2014). The role of paranoia in a dual-process motivational model of conspiracy belief. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139565417.019>.

Capitolul 12. Atribuirea etichetei de fake news ca sursă de structură nespecifică

Jordan R. Axt, Mark J. Landau și Aaron C. Kay

Introducere

În ultimii câțiva ani, afirmația potrivit căreia o informație este „fake news” a devenit un aspect integral al înțelegerii modului în care indivizii percep media și participă la procesul democratic. În general, sintagma „fake news” se referă la știri false, create în scopul de a disemina dezinformare. În acest caz, „fake news” se referă la comunicarea intenționată de informații false în scopul de a induce cititorii în eroare. Un exemplu al acestei forme de fake news este reprezentat de activiștii ruși care au creat reclame pe Facebook cu imagini manipulate în care apăreau celebrități care susțineau că americanii puteau vota de acasă (ceea ce era fals) (Wager, 2017). Un alt exemplu provine de la website-ul 70news, unde a fost un articol, șters între timp, conform căruia Donald Trump câștigase votul popular în alegerile din 2016 (ceea ce era fals), articol care a fost pentru scurt timp unul dintre principalele rezultate oferite de Google News la căutarea „final election results” (rezultatele finale ale alegerilor) (Earl, 2016). În anii de când această formă de fake news a devenit o prezență atât de pregnantă în timpul alegerilor din 2016 din SUA s-au realizat cercetări considerabile asupra motivului pentru care oamenii cred astfel de dezinformări (e.g. Allcott și Gentzkow, 2017; Lazer et al., 2017; Pennycook, Cannon și Rand, 2018, Pennycook și Rand, 2019a), cine sunt cei care au cea mai mare probabilitate de a face asta (e.g. Anthony și Moulding, 2018; Bronstein, Pennycook, Bear, Rand și Cannon,

2018) și asupra modului în care astfel de convingeri pot fi reduse (e.g. Pennycook și Rand, 2019b).

Cu toate acestea, sintagma „fake news“ are și alte sensuri. Această formulă a fost folosită mai ales ca modalitate de a discredita știrile oferite de surse respectabile cu care grupuri sau indivizi nu erau de acord. De exemplu, când ziarul The Courier Journal din Kentucky a publicat un articol conform căruia guvernatorul Matt Bevin și-a cumpărat casa cu un preț de 45% din valoarea de piață de la un susținător politic (Schneider, 2018), Bevin a scris pe Twitter că știrea era „fake news“ și că reporterul care a scris-o era „un om bolnav“ (Manchester, 2017). În acest caz, Bevin susține că The Courier Journal distribuia intenționat informații false în speranța de a-i submina guvernarea, afirmație pe care autorul articolului probabil ar nega-o. Aceste acuzații de „fake news“ sunt folosite pentru a pune la îndoială articole probabil legitime și pentru a atrage atenția asupra posibilității unui efort concentrat și extins prin care instituțiile media intenționează să disemineze informații false în scopul de a submina anumite cauze sau figuri politice.

Cele două sensuri ale sintagmei „fake news“ au origini, funcții și implicații diferite în asemenea măsură, încât, în octombrie 2018, guvernul britanic a decis oficial să nu mai folosească acest termen, dat fiind că a ajuns să desemneze atât de multe procese diferite, de la eroarea umană neintenționată la interferențe străine în procesul democratic. Deși o mare parte a cercetărilor asupra „fake news“ s-au concentrat asupra dezinformării politice, am început să examinăm funcția psihologică a acestei forme alternative de „fake news“ care desemnează modul în care indivizii fac atribuiri de fake news în încercarea de a discredita informațiile din surse respectabile sau legitime.

În acest capitol trecem în revistă modul în care aceste atribuiri de fake news se bazează pe cercetări psihologice anterioare asupra percepției distorsiunilor din media. Apoi, aplicând principiile teoriei controlului compensatoriu (compensatory control theory, CCT; Kay, Gaucher, Napier, Callan și Laurin, 2008), introducem o nouă idee, conform căreia astfel de etichetări drept fake news reprezintă o sursă de „structură nespecifică“ (Landau, Kay și Whitson, 2015), care le permite indivizilor să păstreze o interpretare structurată (deși nefastă) a realității. După care trecem în revistă datele corelaționale și experimentale pe care le-am cules și apoi discutăm unele dintre implicațiile

practice și teoretice ale înțelegerii atribuirilor de etichete fake news ca sursă de structură, inclusiv posibile intervenții care ar putea reduce frecvența unor astfel de acuzații.

Cercetări anterioare asupra percepțiilor distorsionate din media

Distorsiunile de percepție asupra mass-media s-au aflat în centrul a numeroase cercetări psihologice realizate până acum. De exemplu, deși principalele descoperiri din studiul „They saw a game“ (Hastorf și Cantril, 1954) se referă la percepțiile distorsionate ale studenților de la Dartmouth și Princeton, autorii încep expunerea ilustrând modul în care astfel de distorsiuni colorează și articolele din ziarele fiecărei facultăți. După ce Princeton a câștigat un meci deosebit de dur de fotbal american împotriva rivalilor de la Dartmouth, într-un articol din Daily Princetonian scria că atacurile jucătorilor de la Dartmouth erau dovezi clare că jucătorii încercau „intenționat“ să-i rănească pe membrii echipei Princeton. Și invers, un articol al publicației Dartmouth susținea că antrenorul Princeton își instruisese „intenționat“ proprii jucători să-și rănească adversarii. Comparând cele două expuneri, este evident modul în care cititorii își pot forma convingerea că jurnaliștii înșiși pot fi o sursă de distorsionare; desigur, mulți studenți de la Dartmouth care au citit relatările oferite de publicații Princeton (sau viceversa) ar ajunge la concluzia că jurnaliștii presupus obiectivi ilustrau fals evenimente „concrete“ care au avut loc în cadrul meciului. Această suspiciune de distorsionări din partea media pot fi generalizate, astfel încât toate relatările, indiferent de sursă, sunt măcar potențial afectate de partizanatul jurnaliștilor.

Distorsiuni similare de percepție au fost ulterior explorate mai detaliat în contextul „fenomenului de media ostile“ (Vallone, Ross și Lepper, 1985). În această cercetare, autorii au oferit relatări ale unui eveniment internațional important (masacrul din 1982 din Beirut) unor participanți care simpatizau fie

cu israelienii, fie cu arabii. Toți participanții au raportat propriile percepții în ce privește obiectivitatea și corectitudinea echipei de media care a relatat evenimentul. Rezultatele au arătat că participanții pro-israelieni și pro-arabi au văzut relatările media ca fiind la fel de distorsionate împotriva taberei pe care o simpatizau. De exemplu, folosind o scală de la 1 = Distorsiune împotriva Israelului la 9 = Distorsiune în favoarea Israelului, studenții pro-israelieni au considerat înregistrările ca fiind în medie anti-israeliene ($M = 2,9$), în timp ce studenții pro-arabi au considerat aceleași înregistrări ca fiind în medie pro-israeliene ($M = 6,7$). Mai mult, participanții care s-au autoevaluat ca fiind bine informați în ce privește acest subiect au avut convingeri mai puternice că articolele de știri erau distorsionate, nu obiective.

Aceste rezultate pun în lumină mai multe aspecte semnificative ale modului în care indivizii percep media. În primul rând, participanții au considerat în mod consecvent media de știri ca opunându-se, nu susținând, propria perspectivă, indiferent care era perspectiva lor. Acest rezultat este întrucâtva surprinzător, date fiind diferite alte distorsiuni care susțin stima de sine și care sunt implicate în modul în care oamenii percep și își amintesc informațiile; de exemplu, deși este mai probabil să se piardă din memorie informațiile negative (e.g. Zhang, Pan, Li și Guo, 2018) sau ca oamenii să aibă o nevoie generalizată de a prioritiza informațiile care le susțin respectul de sine (e.g. Gaertner, Sedikides și Cai, 2012; Seih, Buhrmester, Lin, Huang și Swann, 2013), procese similare nu apar în modurile în care sunt percepute media de știri. În loc să vadă media ca modalitate de afirmare a propriilor opinii, oamenii par motivați să pună accentul pe aspectele media care nu corespund perspectivei lor asupra lumii.

În plus, faptul că o mai mare cunoaștere autoevaluată a subiectului a fost asociată cu o tendință mai accentuată de a considera media ca fiind distorsionate ilustrează modul în care investirea unei cauze sau a unei chestiuni nu este asociată cu o motivație mai puternică de a se mulțumi cu un adevăr obiectiv, ci cu o motivație mai puternică de a discredita informații potențial conflictuale. Drept urmare, în cazul multor credințe și poate mai ales în cazul convingerilor politice, chiar și oamenii bine informați nu consideră media de știri ca reprezentând un mijloc de a ajunge la o perspectivă nedistorsionată, ci, mai degrabă, o sursă de distorsiune care trebuie descalificată. Într-adevăr, neîncrederea generalizată în media poate

explica parțial polarizarea din ce în ce mai mare a discursului politic (Iyengar și Hahn, 2009). Dat fiind că oamenii privesc media mai degrabă cu suspiciune, pot interacționa cu aceasta în primul rând ca modalitate de a susține ideologiile existente, și nu ca metodă de a ajunge la percepții acurate asupra lumii. Pentru a rămâne relevante, organizațiile media sunt din ce în ce mai presate să răspundă ideologiilor consumatorilor, ceea ce duce la o polarizare din ce în ce mai accentuată în peisajul media.

Nici jurnaliștii înșiși nu sunt scutiți de acest efect al mediei ostile, care constă în faptul că membri ai grupurilor cu perspective opuse consideră același material ca fiind la fel de distorsionat. Într-adevăr, în cazul multor figuri publice, percepțiile asupra presupuselor lor distorsionări sunt atât de răspândite, încât nu pot spera decât să fie văzute ca fiind în mod egal distorsionate, și nu nedistorsionate. De exemplu, James Comey, fostul director FBI care a devenit comentator politic și autor, era mândru de faptul că un alt comentator l-a numit „aservit politic“, dar pur și simplu nu își putea da seama care era tabăra înspre care tindea Comey (Comey, 2018). În mod similar, comentatorii sportivi Joe Buck și Bob Costas au spus amândoi că s-au împăcat cu gândul că este inevitabil să fie considerați părtinitori de fanii ambelor echipe la meciul din finala turneului de baseball World Series (Sandomir, 2009).

Însă, deși există o tradiție îndelungată de cercetări asupra distorsionării în media, este un fenomen mult mai recent faptul că aceste distorsionări se transformă în acuzații de „fake news“, mai precis, că jurnaliștii nu doar că relatează selectiv anumite evenimente, ci răspândesc intenționat informații false pentru a susține o anumită perspectivă. Această acuzație specifică adusă organizațiilor media, conform căreia acestea induc intenționat în eroare, are propriile consecințe în privința modului în care indivizii percep și media de știri, și lumea socială în general.

„Fake news“: din greșeală sau intenționat?

O mare parte din cercetările anterioare asupra modului în care sunt percepute media de știri au descoperit că mulți oameni văd aceste organizații ca surse de distorsiune, în special sub forma distorsiunilor care se opun sau atenuează cauzele pe care cititorii le susțin. Cu toate acestea, acuzațiile de „fake news” mai fac un pas și susțin că media de știri nu doar că sunt distorsionate (i.e. relatează selectiv anumite fapte), ci se înșală (i.e. relatează informații incorecte). În aceste cazuri, reporterii pot să ajungă la prezentarea de informații incorecte din motive neintenționate (e.g. publicarea unei relatări care nu e corectă pentru că nu a fost verificat istoricul oferit de sursă) sau intenționate (e.g. încercarea de a induce cititorii în eroare în speranța de a oferi susținere unei anumite perspective sau ideologii politice).

Într-un sens, aceste atribuiri diferite (i.e. relatări incorecte provocate de greșeli neintenționate sau de inducere intenționată în eroare) servesc aceeași funcție, prin faptul că le permit indivizilor să discrediteze informația care nu corespunde propriei perspective asupra lumii. În consecință, astfel de acuzații de „fake news” pot fi considerate o altă formă de raționament motivat (Kunda, 1990), proces prin care oamenii încearcă să discrediteze informația care le amenință perspectiva asupra lumii. Cu toate acestea, cele două forme de atribuire au implicații profund diferite pentru modul în care operează realitatea socială a individului. A spune că o știre este „fake news” din cauza incompetenței sau a unei erori neintenționate este o modalitate de a o discredita, dar implică perspectiva că lumea socială este aleatorie și lipsită de ordine. O lume în care jurnaliștii răspândesc constant informații false pur și simplu din neglijență este o lume în care apar greșeli grave, în mare imposibil de prevăzut. Pe de altă parte, afirmația că o știre este „fake news” pentru că intenționează să inducă în eroare este reprezentativă pentru o lume răuvoitoare, dar în care există ordine și în care actori cunoscuți se comportă în moduri care pot fi prevăzute și înțelese. Este probabil ca aceste acuzații de „fake news” determinate de inducere intenționată în eroare să fie atrăgătoare în special pentru înțelegerea organizațiilor de știri, care pot fi definite ca grup extins de agenți care se coordonează în funcție de scopuri comune pentru a submina aceleași ținte în mod repetat.

Atribuirea etichetei de „fake news” pe baza inducerii în eroare (și nu a greșelilor omenești) este o opțiune evidentă pentru indivizii care caută să discrediteze sursele de știri cu care nu sunt de acord și în același timp să-și

păstreze o concepție despre lume ca loc ordonat. Această strategie este consecventă rezultatelor cercetărilor anterioare, conform cărora factorii din mediu care amenință controlul personal amplifică tendința indivizilor de a prefera explicații bine structurate ale lumii (i.e. prin structură nespecifică, Landau et al., 2015). În cazul în care resimt amenințări la adresa controlului personal, mulți indivizi pot acorda prioritate unei perspective asupra lumii care are un nivel înalt de structură, chiar dacă această perspectivă implică o realitate în care actori infami se opun activ intereselor individului. Efecte paralele au apărut în cercetări anterioare ca au aplicat diferite forme de amenințări asupra controlului. În cadrul unui studiu, amenințările asupra controlului personal (cum ar fi lectura de materiale despre riscurile încălzirii globale) au dus la o mai mare susținere a concepției că marile companii se străduiesc să distrugă mediul (Rothschild, Landau, Sullivan și Keefer, 2012). În mod similar, participanților cărora li s-a arătat o listă de amenințări asupra cărora aveau foarte puțin control (e.g. expunere la diferite boli) au avut o probabilitate mai mare de a crede că adversarii politici se implicau intenționat în acțiuni neetice (e.g. prin suprimarea voturilor), în comparație cu participanții al căror sentiment de control nu a fost amenințat (Sullivan, Landau și Rothschild, 2010).

În ambele studii, indivizii care au resimțit mai accentuat reducerea controlului personal au încercat să reinstaureze acest control prin susținerea de perspective asupra lumii care o făceau mai predictibilă. Cercetări anterioare sugerează că o modalitate eficientă de a menține percepția structurii este considerarea lumii drept rezultat al acțiunilor intenționate, și nu al unor forțe aleatorii și impersonale. Este semnificativ faptul că această strategie se aplică și în cazul evenimentelor pozitive, și în cazul evenimentelor negative. De exemplu, există o probabilitate mai mare ca vremea bună în ziua nunții să fie atribuită unei divinități binevoitoare, și nu norocului aleatoriu. Totuși, aceste cercetări s-au concentrat în primul rând asupra evenimentelor negative, dat fiind că suntem interesați de modul în care oamenii interpretează informațiile indezirabile în domeniul politic (i.e. știri negative despre figurile politice simpatizate). În aceste cazuri, se poate ca indivizii să încerce să explice evenimentele negative nu prin întâmplări aleatorii, ci prin intențiile răuvoitoare ale celorlalți.

În cazul atribuirilor etichetei de „fake news“, datele anchetelor sugerează că

mulți oameni consideră că se întâmplă mai frecvent ca jurnaliștii să răspândească intenționat relatări false decât să facă greșeli neintenționate. Deși s-ar putea să fie neliniștitor să considerăm că suntem constant expuși la relatări inventate, această perspectivă este destul de răspândită în rândul americanilor. Conform unei anchete recente, 53% dintre democrați, 79% dintre independenți și 92% dintre republicani au spus că au convingerea că sursele tradiționale de știri relatează materiale despre care știu că sunt „false, inventate sau încearcă să inducă în eroare” (Fischer, 2018). În conformitate cu mare parte din cercetările realizate până acum referitor la amenințările asupra controlului, există concepția răspândită că „fake news” sunt rezultatul eforturilor intenționate și coordonate ale anumitor actori. În cercetarea noastră explorăm factorii individuali și de mediu care fac atrăgătoare din punct de vedere psihologic presupuneri precum inducerea intenționată în eroare.

Necesitatea personală de structură și atribuirea etichetei de fake news

O modalitate de studiere a relației dintre atribuirea etichetei de fake news și necesitatea de structură este oferită de examinarea faptului dacă diferențele între indivizi în ce privește nevoia de structură sau dacă manipulările care provoacă experimental o mai mare nevoie de structură sunt asociate cu o disponibilitate mai accentuată de a considera că articole discutabile sunt rezultatul inducerii intenționate în eroare, și nu al erorii umane. În speță, noi considerăm că atribuirea intenției de inducere în eroare materialelor de tip fake news ar trebui să fie asociată cu diferențe la nivelul constructului Necesitate Personală de Structură (Personal Need for Structure, PNS), care evaluează variațiile dorinței ca lumea să fie previzibilă și bine ordonată (Thompson, Naccarato, Parker și Moskowitz, 2001). Cercetări anterioare asupra PNS au arătat că acest construct prezintă validitate divergentă în raport cu constructe posibil conexe, ca autoritarismul, orientarea în situații de

incertitudine și rigiditatea psihologică (Neuberg și Newsom, 1993).

O mare parte dintre cercetările anterioare asupra PNS s-au concentrat asupra unor aspecte cu un nivel relativ redus de elaborare, cum ar fi modul în care PNS are legătură cu procesele social-cognitive de bază. De exemplu, indivizii cu scoruri PNS ridicate au prezentat un nivel mai redus de complexitate atât când li s-a cerut să clasifice stimuli non-sociali (e.g. culori), cât și stimuli sociali (e.g. indicatori ai vârstei înaintate) și au prezentat o încredere mai mare în stereotipuri de gen în cadrul unei sarcini centrate pe prima impresie (Neuberg și Newsom, 1993). Cu toate acestea, cercetări mai recente au abordat înțelegerea constructului ca fenomen motivat afectiv, ceea ce înseamnă că indivizii cu scoruri PNS ridicate văd o lume structurată nu doar ca preferință personală, ci ca sursă de sens. Din această perspectivă, PNS nu este pur și simplu un stil cognitiv sau o tendință imparțială de a gândi în moduri simple și clare. Mai degrabă descrie diferențe individuale în privința gradului de structură pe care indivizii își doresc ca lumea să îl aibă. Această perspectivă motivațională duce la o ipoteză testabilă: când indivizii cu scoruri PNS ridicate sunt în situații care amenință structura — cum ar fi complexitatea, inconsecvența sau ambiguitatea — vor reacționa prin eforturi cognitive exagerate de a restaura structura. Nu ne-am aștepta ca această reacție compensatorie să structureze amenințările dacă PNS ar fi un stil cognitiv „rece“, fără motivații afective.

De exemplu, doar indivizii cu scoruri PNS mari au raportat o senzație mai accentuată de sens al vieții în urma unei amenințări mortale și asociații paralele cu viața plină de sens au fost identificate în rândul persoanelor care aveau și scoruri PNS ridicate, și gânduri referitoare la moarte (death-thought accessibility) (Vess, Routledge, Landau și Arndt, 2009). În plus, doar indivizii cu scoruri PNS ridicate au prezentat o nevoie mai accentuată de a acorda prioritate tradiției ca reacție la o amenințare legată de terorism (Routledge, Juhl și Vess, 2010). În sfârșit, persoanele cu scoruri PNS mari prezintă niveluri mai reduse de anxietate legată de moarte în urma perceperii unei amenințări mortale (Routledge, Juhl și Vess, 2013). Împreună, aceste cercetări mai recente sugerează că PNS este un fenomen parțial motivat afectiv; oamenii care au în mod natural o nevoie personală de structură (PNS) accentuată (sau cărora li se induce această nevoie) nu preferă pur și simplu o lume în mai mare măsură structurată, ci folosesc această preferință a unei

lumi structurate ca sursă de sens (Swanson și Landau, în curs de publicare).

Pe baza acestor cercetări de până acum, noi am dezvoltat mai multe ipoteze referitoare la relația dintre PNS și atribuirea intenției de inducere în eroare materialelor de tip fake news. În primul rând, la nivel corelațional, PNS ar trebui să fie asociată cu credința că „fake news” (i.e. știrile discutabile) sunt mai degrabă rezultatul unei amăgiri intenționate din partea jurnaliștilor, și nu o serie de greșeli omenești neintenționate, dar nefaste. În al doilea rând, această asociere între PNS și atribuirea intenției de inducere în eroare materialelor de tip fake news ar trebui să se aplice și în cazul știrilor care susțin ideologia persoanelor (i.e. când indivizii citesc știri care prezintă într-o lumină negativă un personaj politic antipatizat), și în cazul știrilor care nu susțin ideologia persoanelor (i.e. când indivizii citesc știri care prezintă într-o lumină negativă personaje politice simpatizate). Cu alte cuvinte, întrucât convingerea că media de știri induce intenționat în eroare este legată de o dorință mai mare de a vedea lumea ca loc ordonat, acest efect ar trebui să apară și când știrile corespund ideologiei politice a persoanei, și când nu corespund. O astfel de analiză poate elimina presupunerea inducerii intenționate în eroare ca fiind pur și simplu o formă de favoritism în cadrul grupului (i.e. retractarea știrilor negative despre candidați simpatizați este o distorsiune intenționată, dar retractarea știrilor negative despre candidați antipatizați este o greșeală omenească). În loc de asta, presupunerea existenței unei intenții de inducere în eroare ar reflecta mai degrabă dorința de a menține percepția structurii.

În sfârșit, pentru a oferi dovezi cauzale ale relației dintre PNS și atribuirea etichetei de fake news, indivizii repartizați aleatoriu într-un grup în care sentimentul de control personal este amenințat ar trebui să prezinte o tendință mai accentuată de a atribui erorile din știri unei intenții de inducere în eroare, și nu unor simple greșeli provocate de incompetență. Noi considerăm că o astfel de cercetare are potențialul de a duce la progres în ce privește înțelegerea proceselor psihologice care stau la baza atractivității din ce în ce mai mare a invocării faptului că știrile sunt „fake news”.

Studierea invocării fake news

O provocare inevitabilă în studiul invocării fake news este natura în mod fundamental subiectivă a convingerii. Pentru mulți oameni nu există nicio diferență între instigatorii ruși care răspândesc intenționat informații false — cum ar fi articole care susțin că este posibil votul de acasă sau că Donald Trump a câștigat votul popular — și știri discutabile care provin de la organizații respectabile, cum ar fi relatarea din The Courier Journal a prețului suspect de mic pe care guvernatorul Kentucky, Matt Bevin, l-a plătit pentru locuință. Acuzații similare de fake news au fost legate de cele mai notabile nume din media, ca The New York Times, The Washington Post și CNN. Cu toate acestea, legitimitatea percepută nu este suficientă pentru a convinge cititorii sceptici că anumite știri nu răspândesc deliberat informații false.

Pentru a evita această problemă, noi considerăm că este productiv să ne concentrăm asupra cazurilor în care a fost necesar ca organizații de știri să corecteze sau să retracteze articole. Aceste retractări și corectări oferă astfel un exemplu clar de situație în care reporterii au făcut afirmații greșite, ceea ce îi obligă pe cititori să evalueze motivul pentru care au apărut astfel de erori — fie ca efect al unei greșeli omenești, fie ca inducere intenționată în eroare. De exemplu, în mai 2017, Fox News a publicat online un articol despre uciderea în 2016 a lui Seth Rich (Grynbaum, 2018), fost angajat al Comitetului Democratic Național (Democratic National Committee, DNC). Printre alte detalii, articolul specula că moartea lui Rich era cumva legată de scurgerea de informații prin e-mailuri de la persoane proeminente în DNC, ca John Podesta. În speță, articolul sugera faptul că s-ar fi putut ca Rich să fi fost ucis de oameni aflați în slujba DNC ca răzbunare pentru posibilul rol pe care l-a jucat în scurgerea de informații.

La scurt timp după apariția articolului, mai multe organizații de știri au atacat Fox News pentru publicarea acestui articol (e.g. Pilkington, 2017). Aceste critici subliniau evidenta lipsă de dovezi care să facă legătura între persoane asociate DNC și uciderea lui Rich. În cele din urmă, Fox News a retractat articolul și a eliminat acest conținut din website pentru că nu se ridica la „înaltul standard de calitate redacțională” al site-ului (Statement on coverage of Seth Rich, 2017). Pentru apărătorii Fox News, articolul retractat nu era decât rezultatul unei erori omenești dar nefaste apărute în procesul jurnalistic

firesc, de exemplu cauzate de neconfirmarea relatărilor unor surse anonime. Pentru alții, articolul era un exemplu de răspândire de informații false de către organizații partinitoare de știri, în speranța de a susține o anumită cauză politică, precum punerea DNC într-o lumină negativă. Publicul a avut de-a face cu aceeași problemă: era articolul retractat un exemplu de situație în care jurnaliștii au făcut erori firești în procesul de relatare sau pur și simplu Fox News a fost prinsă într-un efort mai amplu de înșelare intenționată a audienței? Articolul despre Seth Rich, precum și alte articole retractate, obligă membrii audienței să aleagă între explicația unei erori mai puțin nocive, dar aleatoare, și explicația unei campanii coordonate, rău-voitoare, dar structurate.

În studiile noastre inițiale asupra acestei chestiuni (Axt, Landau și Kay, în curs de apariție) le-am oferit participanților descrieri ale unor știri care au fost retractate sau corectate. În cadrul unui studiu, participanții nu au văzut decât articole consecvente din punct de vedere ideologic, prezentate în mai multe viniete, în general articole retractate care prezentau într-o lumină negativă membrii partidului politic pe care participanții îl susțineau (e.g. participanții democrați citeau despre articolul privitor la Seth Rich al Fox News, în timp ce participanții republicani citeau despre articolul publicat de ABC News, conform căruia Michael Flynn era pregătit să depună mărturie că Donald Trump îl instruisese să colaboreze cu oficiali ruși pentru influențarea alegerilor prezidențiale din 2016). După ce citeau fiecare vinietă, participanții răspundeau la mai mulți itemi prin care se evalua gradul în care ei considerau că articolul retractat era fie rezultatul unor erori omenești, fie al dorinței de a răspândi informații false. De exemplu, participanții raportau dacă credeau că e mai probabil ca eroarea să fie cauzată de neatenție sau de un atac intenționat din partea jurnalistului. Răspunsurile au fost notate pe o scală Likert de șase puncte, astfel încât participanții să poată indica gradul de încredere în răspunsurile pe care le ofereau („Extrem de probabil să fie o greșeală omenească“, „Probabilitate moderată de a fi o eroare omenească“ etc.).

Pe baza acestor itemi am dezvoltat apoi un instrument caracterizat de consecvență internă care evalua gradul în care participanții credeau că articolele retrase erau rezultatul intenției de inducere în eroare. Am folosit această metodă de evaluare și am examinat gradul în care convingerea că știrile eronate erau provocate de intenția de inducere în eroare corela sau era

cauzată de nevoia personală de structură.

La un nivel descriptiv, rezultatele au fost compatibile cu cele din ancheta menționată anterior, conform căreia mulți democrați și republicani consideră că organizațiile media tradiționale răspândesc intenționat informații false în speranța de a-și îndeplini propriile obiective politice. De exemplu, în primul nostru studiu, 72,7% dintre democrați credeau că articolul despre Seth Rich avea la bază intenția de inducere în eroare și 84,9% credeau că la fel stăteau lucrurile și în privința relatării „Pizzagate“ a companiei media Breitbart, care indica posibilitatea existenței unei rețele de proxenetism în care erau implicate mai multe persoane proeminente din Partidul Democrat. Dintre republicani, 53,5% credeau că articolul publicat de ABC News despre Michael Flynn a fost creat de jurnaliști care aveau intenția clară de a induce în eroare și 55,4% au oferit aceeași explicație pentru un articol al „MSNBC“ care susținea că Wikileaks le-a oferit unor membri proeminenți ai Partidului Republican acces special la e-mailuri DNC care mai târziu au fost postate online. Deși democrații păreau să aibă o rată mai mare de atribuire a intenției de inducere în eroare, diferențele din materialele studiului fac dificilă comparația ratelor între participanții democrați și republicani. În sfârșit, deși independenții nu au fost incluși în analiza primară, rezultatele acestora au arătat că percepția intenției de inducere în eroare nu este limitată la partizani; în același studiu, 58,3% dintre independenți credeau că articolul Fox News despre Seth Rich era o încercare de inducere în eroare și 40% credeau același lucru în privința relatării ABC News despre Michael Flynn.

Cel mai important aspect este că am observat în toate studiile o relație consecventă și pozitivă între valorile PNS măsurate (Thompson et al., 2001; exemple de itemi: „Mă simt inconfortabil dacă, într-o anumită situație, regulile nu sunt clare“, „Îmi place să am un model de viață clar și structurat“) și o convingere accentuată că organizațiile de știri încearcă activ să-și înșele audiența. În studii realizate mai târziu am descoperit că această relație pozitivă se menține și în cazul în care persoanele află despre știri eronate care sunt în conflict sau care sunt în acord cu convingerile lor politice, și se mențin, de asemenea, dacă sunt controlate diferențele individuale de intensitate a credințelor ideologice. Aceste descoperiri sugerează că asociația dintre PNS și presupunerea de inducere intenționată în eroare reflectă mai mult decât doar o dorință de susținere a propriului grup politic sau alte forme

de distorsiune ideologică; se pare, mai degrabă, că PNS prezice presupunerea de inducere intenționată în eroare în cazul știrilor care prezintă negativ și membrii propriului grup, și membrii altui grup politic.

Am încercat, de asemenea, să găsim dovezi cauzale ale influenței PNS asupra presupunerii de fake news prin aplicarea unui design experimental.

Cercetările de până acum au descoperit că cele mai eficiente manipulări ale sentimentului de structură sunt oferite de amenințările asupra controlului personal (vezi Landau et al., 2015). Am repartizat participanții în mod aleatoriu și le-am cerut să reflecteze asupra unor aspecte ale vieții lor care erau fie relativ controlabile, fie incontroleabile (vezi Kay et al., 2008 pentru aplicări anterioare ale acestei manipulări). Rezultatele au arătat că participanții care au reflectat la aspecte incontroleabile ale vieții au raportat o convingere mai accentuată că știrile eronate erau provocate de intenția de inducere în eroare, prin comparație cu greșeli omenești. Am observat, de asemenea, că reacția la amenințare printr-o atribuire mai accentuată a distorsiunii în prezentarea știrilor contribuie la menținerea sentimentului că lumea este structurată.

Așadar, rezumând, cercetările noastre subliniază și o relație corelațională, și o relație cauzală între PNS și convingerea că media încearcă să inducă în eroare prin știri de tip „fake news”. Prin aceste studii, proiectul nostru ilustrează una dintre potențialele funcții ale fake news și oferă o bază pentru considerații mai generale în privința modului în care fake news pot fi mai bine înțelese și chiar reduse.

Progresul tehnologic și reducerea influenței fake news

Considerăm că cercetările noastre demonstrează o legătură între PNS și presupunerea intenției de inducere în eroare de către media de știri (i.e. acuzația de „fake news”). Deși descoperirea acestei asociații este importantă,

ea duce în mod firesc la întrebări mai practice, referitoare la modul în care putem aplica descoperirile de până acum referitoare la PNS în cazul atenuării problemei fake news. Din fericire, studiile realizate până în prezent asupra necesității de structură nespecifică sugerează mai multe modalități prin care poate fi atenuată presupunerea intenției de inducere în eroare de către media de știri. Poate cea mai directă implicație vine din cercetarea asupra beneficiilor oferite de sentimentul de structură, similar manipulării pe care am aplicat-o când le-am cerut participanților să reflecteze la o perioadă din viața lor în care au simțit că dețin controlul. La fel cum privarea participanților de sentimentul structurii duce la o nevoie mai accentuată de a atribui ordine sistemelor externe (e.g. Kay, Shepherd, Blatz, Chua și Galinsky, 2010; Wang, Whitson și Menon, 2012), faptul de a le oferi o senzație de control sau de structură atenuază această necesitate. De exemplu, participanții care au scris despre o perioadă din viața lor în care au avut un nivel ridicat de control au prezentat mai târziu percepții reduse ale ierarhiei într-o interacțiune socială ambiguă și au raportat o preferință mai redusă pentru ierarhie (i.e. sistem de structură și ordine) când evaluau opțiuni de locuri de muncă (Friesen, Kay, Eibach și Galinsky, 2014).

Aceste rezultate și propriile noastre cercetări arată clar că afirmarea structurii oferă o modalitate viabilă de reducere a atribuirii etichetei de fake news; simplu spus, punerea pe seama jurnaliștilor a unei intenții răuvoitoare, dar previzibile, ar trebui să fie mai puțin atractivă pentru indivizii care au un sentiment mai accentuat de control personal. E posibil ca aplicații practice ale unor astfel de rezultate să necesite creativitate; nu este foarte clar modul în care manipulările experimentale ale controlului pot fi încorporate firesc în modurile în care oamenii primesc știri. O posibilitate este oferită de social media și de noile forme de comunicare. Deși social media sunt poate cea mai mare sursă de fake news — un studiu recent a descoperit că 8,5% dintre utilizatorii de Facebook au distribuit cel puțin un articol care poate fi considerat dezinformare politică (Guess, Nagler și Tucker, 2019) —, este posibil și să reprezinte cea mai mare oportunitate pentru ca oamenii să-și recâștige senzația de control personal. Interacțiunea cu alte persoane cunoscute, împărtășirea sentimentelor referitoare la propria viață sau organizarea propriului flux de informații (e.g. prin eliminarea sau adăugarea de prieteni) pot fi câteva modalități prin care oamenii pot să resimtă că au întrucâtva control asupra propriei vieți, iar aceste posibilități sunt din ce în ce

mai comune pe platformele de social media.

Una dintre predicțiile care se pot face pe baza cercetărilor din prezent este că participanții cărora li s-a cerut să interacționeze cu social media într-o modalitate mai directă, controlată, pot chiar să prezinte o nevoie mai mică de a atribui motivații de inducere intenționată în eroare din partea organizațiilor de știri. Cu alte cuvinte, instruirea în privința evaluării atente a informației pe care o văd pe platformele de social media ar putea duce la un nivel mai înalt de percepere a structurii în lume. Mai mult, beneficiile afirmării stucturii probabil nu se limitează la acuzațiile de fake news, dat fiind că alte cercetări sugerează că sentimente accentuate de control pot reduce intensitatea anxietății (Tullet, Kay și Inzlicht, 2014) sau pot facilita urmărirea obiectivelor (Kay, Laurin, Fitzsimons și Landau, 2014). Inducerea sentimentelor de control are potențialul de a atenua și alte consecințe negative, nu doar încrederea în fake news. În același timp, se poate ca oamenii să resimtă și consecințe benefice, și consecințe nocive, întrucât platformele social media sunt modelate în funcție de preferințe individuale; astfel de modificări pot crea percepții ale unei lumi mai structurate și pot în același timp să întărească bule de partizanat, eliminând ocazia de a vedea informații nealiniate propriei perspective asupra lumii.

O altă sursă posibilă de restaurare a sentimentelor de control poate proveni nu din schimbarea intensității senzației individuale de control, ci din modificarea percepțiilor asupra înseși mediei de știri. Neîncrederea și percepțiile eronate în privința jurnalismului sunt răspândite. De exemplu, un studiu recent (Columbia Journalism Review, 2019) a descoperit că mai puțin de 25% dintre democrați și 15% dintre republicani raportează că au „mare încredere” în media de știri. De fapt, dintre alte instituții, cum ar fi congresul, armata și curtea supremă, media a primit cele mai multe răspunsuri că participanții „nu au aproape deloc încredere”. O mare parte din această neîncredere în media de știri poate fi provocată de convingerile eronate în privința modului în care operează jurnalismul. De exemplu, în același studiu, 60% dintre respondenți au spus că au convingerea că jurnaliștii își plătesc sursele și un procent similar credea că jurnaliștii scriu articolele înainte de a afla faptele.

Data fiind neîncrederea generală în media de știri și percepțiile eronate în privința modului în care operează jurnalismul, acuzațiile de fake news pot fi

atenuate și prin amplificarea transparenței procesului jurnalistic. Adică, un motiv pentru care oamenii ar putea avea senzația că e întemeiată atribuirea unei intenții de inducere în eroare organizațiilor media este că le lipsește înțelegerea procesului jurnalistic, de la confirmarea faptelor din mai multe surse la editare și verificarea faptelor. Este plauzibil ca o mai mare înțelegere a acestor procese (și a ratei relativ mici de erori pe care le produc) să reducă neîncrederea cititorilor în organizațiile de știri și ideea că acestea servesc agendele jurnaliștilor. Potențiale schimbări mici, cum ar fi ca sub semnătura autorului să apară persoanele care s-au ocupat de verificarea faptelor și editorii sau trecerea în titlu a numărului de surse care confirmă relatarea sunt lucruri care pot amplifica încrederea în media și care pot reduce tendința de a presupune că jurnaliștii încearcă să inducă în eroare. Și progresul tehnologic în ce privește consumul media ar putea face posibile niște schimbări mai sofisticate; de exemplu, cititorilor online li s-ar putea arăta documente relevante pentru articolul pe care îl citesc. Se poate ca aceste intervenții să nu schimbe nivelurile individuale de PNS, dar pot redirecționa dorința de structură spre alte surse (e.g. guvern sau religie; Kay et al., 2008; Kay Gaucher, McGregor și Nash, 2010) și pot astfel reduce posibilitatea ca structura să fie restaurată prin punerea unor intenții răuvoitoare în seama jurnaliștilor.

Noi anticipăm mai multe posibilități pentru viitoarele direcții în această sferă de cercetare. O aplicație clară este reprezentată de alegerile prezidențiale din 2020 din SUA, unde problema „fake news” va fi probabil iarăși foarte relevantă. De exemplu, alegerile din 2020 fac posibilă colectarea longitudinală de date, cercetătorii având posibilitatea de a identifica schimbările individuale în privința necesității personale de structură și pot evalua modul în care aceste schimbări se traduc în neîncredere în media de știri. În mod similar, deși dovezile noastre experimentale sugerează că schimbările la nivelul sentimentului individual de structură duc la o dorință mai accentuată de a crede în intenția de inducere în eroare, este posibil să existe o relație bidirecțională, neîncrederea din ce în ce mai mare în organizațiile de știri putând duce la schimbări care apar mai târziu la nivelul necesității de structură. În sfârșit, anticipăm că această cercetare poate fi aplicată în cazul identificării persoanelor care sunt mai susceptibile să presupună că e vorba de „fake news”; de exemplu, cercetări anterioare au descoperit că e posibil ca indivizii cu statut social scăzut (e.g. Fiori, Brown,

Cortina și Antonucci, 2006) să aibă un nivel redus cronic de control personal, și, în consecință, tind să considere deosebit de atrăgătoare atribuirea intenției de a induce în eroare organizațiilor media.

Concluzie

Tendința din ce în ce mai accentuată de a considera știrile (adesea difuzate de organizații legitime) ca fiind „fake news” — presupunând că jurnaliștii intenționează să inducă publicul în eroare prin răspândirea de informații false — reprezintă o amenințare serioasă nu doar la adresa încrederii în media, ci și în procesul democratic în general. Pe baza cercetărilor anterioare asupra teoriei controlului compensator și a necesității personale de structură, noi considerăm că multe atribuiri de intenții de inducere în eroare sunt atractive pentru că oferă o modalitate prin care indivizii pot găsi în mediul lor surse de structură nespecifică. O mai bună înțelegere a proceselor psihologice care determină astfel de acuzații de fake news va accelera progresul teoretic și practic în privința modului în care această atribuire poate fi înțeleasă și, posibil, redusă.

Referințe

Allcott, H. și Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31, 211–236.

Anthony, A. și Moulding, R. (2018). Breaking the news: Belief in fake news and conspiracist beliefs. *Australian Journal of Psychology*. Advance online publication. <https://doi.org/10.1111/ajpy.12233>

Axt, J.R., Landau, M.J. și Kay, A.C. (2020). The psychological appeal of fake news attributions. *Psychological Science*.
<https://doi.org/10.1177/0956797620922785>

Bronstein, M.V., Pennycook, G., Bear, A., Rand, D.G. și Cannon, T.D. (2019). Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *Journal of Applied Research in Memory and Cognition*. Advance online publication.
<https://doi.org/10.1016/j.jarmac.2018.09.005>

Columbia Journalism Review. (2019). Poll: How does the public think journalism happens? Accesat la www.cjr.org/

Comey, J. (30 septembrie 2018). James Comey: The F.B.I can do this. Accesat la www.nytimes.com/

Earl, J. (14 noiembrie 2016). Google's top search result for „final election numbers“ leads to fake news site. Accesat la www.cbsnews.com

Fiori, K.L., Brown, E.E., Cortina, K.S. și Antonucci, T.C. (2006). Locus of control as a mediator of the relationship between religiosity and life satisfaction: Age, race, and gender differences. *Mental Health, Religion and Culture*, 9, 239–263.

Fischer, S. (27 iunie 2018). 92% of Republicans think media intentionally reports fake news. Accesat la www.axios.com/

Friesen, J.P., Kay, A.C., Eibach, R.P., și Galinsky, A.D. (2014). Seeking structure in social organization: Compensatory control and the psychological advantages of hierarchy. *Journal of Personality and Social Psychology*, 106, 590–609.

Gaertner, L., Sedikides, C. și Cai, H. (2012). Wanting to be great and better but not average: On the pancultural desire for self-enhancing and self-improving feedback. *Journal of Cross-Cultural Psychology*, 43, 521–526.

Grynbaum, M.M. (13 martie 2018). Family of Seth Rich sues Fox News over retracted article. *The New York Times*. Accesat la

www.nytimes.com/2018/03/13/business/fox-news-seth-rich-lawsuit.html

Guess, A., Nagler, J., și Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, 5, eaau4586.

Hastorf, A.H. și Cantril, H. (1954). They saw a game: A case study. *The Journal of Abnormal and Social Psychology*, 49, 129–134.

Iyengar, S. și Hahn, K.S. (2009). Red media, blue media: Evidence of ideological selectivity in media use. *Journal of Communication*, 59, 19–39.

Kay, A.C., Gaucher, D., McGregor, I. și Nash, K. (2010). Religious belief as compensatory control. *Personality and Social Psychology Review*, 14, 37–48.

Kay, A.C., Gaucher, D., Napier, J.L., Callan, M.J. și Laurin, K. (2008). God and the government: Testing a compensatory control mechanism for the support of external systems. *Journal of Personality and Social Psychology*, 95, 18–35.

Kay, A.C., Laurin, K., Fitzsimons, G.M. și Landau, M.J. (2014). A functional basis for structure-seeking: Exposure to structure promotes willingness to engage in motivated action. *Journal of Experimental Psychology: General*, 143, 486–491.

Kay, A.C., Shepherd, S., Blatz, C.W., Chua, S.N. și Galinsky, A.D. (2010). For God (or) country: The hydraulic relation between government instability and belief in religious sources of control. *Journal of Personality and Social Psychology*, 99, 725–739.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480–498.

Landau, M.J., Kay, A.C. și Whitson, J.A. (2015). Compensatory control and the appeal of a structured world. *Psychological Bulletin*, 141, 694–722.

Lazer, D., Baum, M., Grinberg, N., Friedland, L., Joseph, K., Hobbs, W. și Mattsson, C. (2017). Combating fake news: An agenda for research and

action. Accesat la <https://shorensteincenter.org/combating-fake-news-agenda-for-research/>

Manchester, J. (28 mai 2017). Kentucky Governor calls reporter a „sick man“. Accesat la <https://thehill.com/homenews/news/335487-kentucky-governor-calls-reporter-a-sick-man>

Murphy, M. (23 octombrie 2018). Government bans phrase „fake news“. Accesat la www.telegraph.co.uk/

Neuberg, S.L. și Newsom, J.T. (1993). Personal need for structure: Individual differences in the desire for simpler structure. *Journal of Personality and Social Psychology*, 65, 113–131.

Pennycook, G., Cannon, T.D. și Rand, D.G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology General*, 147, 1865–1880.

Pennycook, G. și Rand, D.G. (2019a). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*. Advance online publication. <https://doi.org/10.1016/j.cognition.2018.06.011>

Pennycook, G. și Rand, D.G. (2019b). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences*, 201806781.

Pilkington, E. (7 august 2017). The strange case of Fox News, Trump and the death of young Democrat Seth Rich. Accesat la www.theguardian.com/

Rothschild, Z.K., Landau, M.J., Sullivan, D. și Keefer, L.A. (2012). A dual-motive model of scapegoating: Displacing blame to reduce guilt or increase control. *Journal of Personality and Social Psychology*, 102, 1148–1163.

Routledge, C., Juhl, J. și Vess, M. (2010). Divergent reactions to the terror of terrorism: Personal need for structure moderates the effects of terrorism salience on worldview-related attitudinal rigidity. *Basic and Applied Social Psychology*, 32, 243–249.

Routledge, C., Juhl, J. și Vess, M. (2013). Mortality salience increases death-anxiety for individuals low in personal need for structure. *Motivation and Emotion*, 37, 303–307.

Sandomir, R. (2 noiembrie 2009). Some fans see enemies behind every microphone. Accesat la www.nytimes.com/

Schneider, G. (27 aprilie 2018). Matt Bevin's Anchorage home now valued at almost twice what he paid in controversial deal. Accesat la www.courier-journal.com/

Seih, Y.T., Buhrmester, M.D., Lin, Y.C., Huang, C.L. și Swann, W.B., Jr. (2013). Do people want to be flattered or understood? The cross-cultural universality of self-verification. *Journal of Experimental Social Psychology*, 49, 169–172.

Statement on coverage of Seth Rich murder investigation. (23 mai 2017). Accesat la www.foxnews.com/

Sullivan, D., Landau, M.J. și Rothschild, Z.K. (2010). An existential function of enemyship: Evidence that people attribute influence to personal and political enemies to compensate for threats to control. *Journal of Personality and Social Psychology*, 98, 434–449.

Swanson, T.J. și Landau, M.J. (în curs de tipărire). Terror management motivation fuels structureseeking. În C. Routledge și M. Vess (coord.), *The handbook of terror management theory*. Elsevier.

Thompson, M.M., Naccarato, M.E., Parker, K.C. și Moskowitz, G.B. (2001). The personal need for structure and personal fear of invalidity measures: Historical perspectives, current applications, and future directions. În G.B. Moskowitz (coord.), *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* (pp. 19–39). Mahwah, NJ: Erlbaum.

Tullett, A.M., Kay, A.C. și Inzlicht, M. (2014). Randomness increases self-reported anxiety and neurophysiological correlates of performance monitoring. *Social Cognitive and Affective Neuroscience*, 10, 628–635.

Vallone, R.P., Ross, L. și Lepper, M.R. (1985). The hostile media phenomenon: Biased perception and perceptions of media bias in coverage of the Beirut massacre. *Journal of Personality and Social Psychology*, 49, 577–585.

Vess, M., Routledge, C., Landau, M.J. și Arndt, J. (2009). The dynamics of death and meaning: The effects of death-relevant cognitions and personal need for structure on perceptions of meaning in life. *Journal of Personality and Social Psychology*, 97, 728–744.

Wagner, K. (31 octombrie 2017). These are some of the tweets and Facebook ads Russia used to try and influence the 2016 presidential election. Accesat la www.recode.net/

Wang, C.S., Whitson, J.A. și Menon, T. (2012). Culture, control, and illusory pattern perception. *Social Psychological and Personality Science*, 3, 630–638.

Zhang, Y., Pan, Z., Li, K. și Guo, Y. (2018). Self-serving bias in memories: Selectively forgetting the connection between negative information and the self. *Experimental Psychology*, 65, 236–244.

***** Distorsiunea substituției (attribute substitution sau substitution bias) desemnează procesul psihologic prin care, în cadrul unui raționament complex, variabila evaluată (target attribute) este înlocuită cu o alta, mai ușor de accesat. De exemplu, stereotipurile: nivelul de inteligență, dificil de evaluat, este presupus pe baza culorii pielii (prejudecata rasială înlocuind efortul reflexiv). (N.t.)

***** Fenomen studiat de psihologia socială și de științele cognitive, prin care oamenii folosesc un raționament distorsionat emoțional pentru a produce justificări sau a lua decizii dorite, în detrimentul celor care reflectă în mod adecvat realitatea. (N.r.)

***** Problemă de software preconizată la sfârșitul anilor 1990 ca reprezentând un risc major pentru funcționarea sistemelor informatice — majoritatea programelor foloseau afișarea ultimelor două cifre ale anului, ceea ce ar fi însemnat că, la trecerea în noul mileniu, data afișată de computere ar fi trecut de la '99 la '00, adică înapoi la 1900. Au existat numeroase ipoteze referitoare la posibilele reacții în lanț provocate de această eroare logică, de la o criză economică la prăbușirea întregii infrastructuri. (N.t.)

Mulțumiri

Numeroși oameni ne-au ajutat să alcătuim acest volum. Suntem extrem de recunoscători pentru sprijinul lor. Le mulțumim tuturor celor care au participat la procesul de evaluare științifică a articolelor, pentru că au făcut asta și pentru că le-au oferit coordonatorilor și autorilor feedback și sugestii de o utilitate inestimabilă. Către Saskia Bollin se îndreaptă recunoștința noastră pentru asistența ei neprețuită în toate etapele proiectului. Le mulțumim tuturor celor de la Routledge care au participat la acest proces încă de la prima propunere de publicare a cărții, în special lui Eleanor Taylor. Și mulțumim Fundației Naționale pentru Știință din Elveția pentru sprijinul său generos (grantul #IZSEZ0_180211 și grantul #10BP12_193570).

Tabelul 4.6. Citări, cititori și tweet-uri ale celor mai citate, citite și distribuite articole retractate

	Anul publicării	Anul retractări
Cele mai multe citări pre și post-retractare	2011	2014
Cele mai multe citări pre-retractare	2010	2016
Cele mai multe citări post-retractare	2012	2013
Cele mai multe citiri și tweet-uri (A și P)	2013	2014

Notă: A de la Altmetric.com și P de la PlumX

Figura 5.1. Fontul și identificarea informațiilor menite a induce în eroare

Font	Procentul răspunsurilor
Câte animale din fiecare specie a luat Moise pe arcă?	88%
Câte animale din fiecare specie a luat Moise pe arcă?	53%

Sursă: adaptată de la Song și Schwarz (2008), Experimentul 1.